# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-097241
(43)Date of publication of application : 08.04.1997

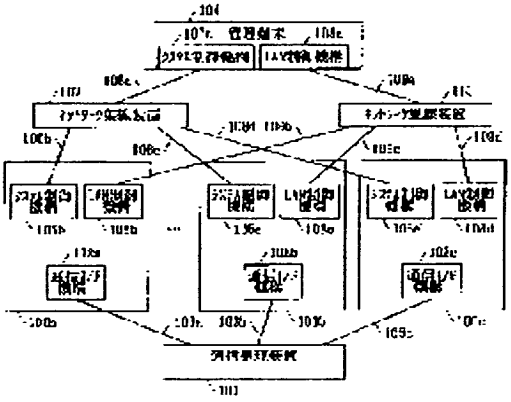
(51)Int.Cl. G06F 15/16
G06F 1/26

(21)Application number : 07-251427
(22)Date of filing : 28.09.1995

(71)Applicant : HITACHI LTD
(72)Inventor : MATSUSHITA SUKEYUKI
UGAJIN ATSUSHI

## (54) MANAGEMENT EQUIPMENT FOR PARALLEL COMPUTER SYSTEM

(57)Abstract:
PROBLEM TO BE SOLVED: To integrally perform the operation management for plural nodes composing a parallel computer system by a managing terminal equipment by transmitting a system control command from the managing terminal equipment to the plural sub-processors of plural nodes.
SOLUTION: The system control interface of the management equipment for a parallel computer system is the interface realized by performing the mutual connection of the system control mechanism 105a on the side of a control terminal equipment 104 and the system control mechanisms 105b is 105d on the side of nodes 100a to 100c by using a communication cable 106 such as an Ethernet, etc., and a network line concentration device 107 such as a multiport repeater, etc. The system control interface transmits the system control commands managing plural main processors of the plural nodes 100a to 100c from the managing terminal equipment 104 to the plural subprocessors of the plural nodes 100a to 100c.



LEGAL STATUS

| | |
|---|---|
| [Date of request for examination] | 04.03.1999 |
| [Date of sending the examiner's decision of rejection] | 03.10.2000 |
| [Kind of final disposal of application other than the examiner's decision of rejection or application converted registration] | |
| [Date of final disposal for application] | |
| [Patent number] | 3163237 |
| [Date of registration] | 23.02.2001 |
| [Number of appeal against examiner's decision of rejection] | 2000-17559 |

[Date of requesting appeal against examiner's decision of rejection] 02.11.2000

[Date of extinction of right]

(19)日本国特許庁（ＪＰ）　(12) **公開特許公報（A）**　(11)特許出願公開番号

**特開平9－97241**

(43)公開日　平成９年(1997)４月８日

| (51)Int.Cl.[6] | 識別記号 | 庁内整理番号 | ＦＩ | | 技術表示箇所 |
|---|---|---|---|---|---|
| Ｇ０６Ｆ　15/16 | | | Ｇ０６Ｆ　15/16 | ４２０Ｃ | |
| | | | | Ｅ | |
| 1/26 | | | 1/00 | ３３４Ｈ | |

審査請求　未請求　請求項の数10　ＯＬ　（全 34 頁）

(21)出願番号　特願平7－251427

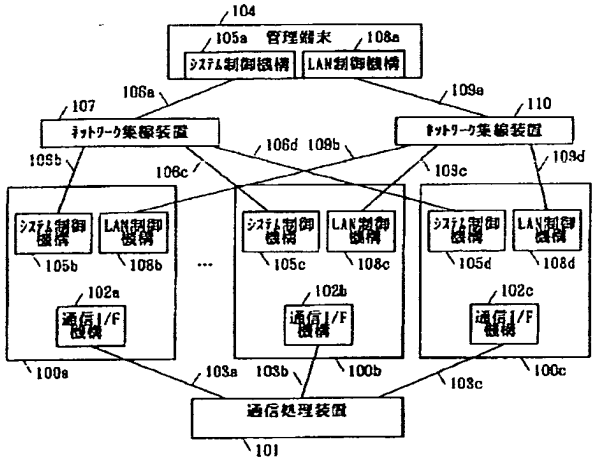
(22)出願日　平成７年(1995)９月28日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地
(72)発明者　松下　祐之
神奈川県海老名市下今泉810番地　株式会社日立製作所オフィスシステム事業部内
(72)発明者　宇賀神　敦
神奈川県海老名市下今泉810番地　株式会社日立製作所オフィスシステム事業部内
(74)代理人　弁理士　秋田　収喜

(54)【発明の名称】　並列計算機システムの管理装置

(57)【要約】
【課題】　並列計算機システムを構成する複数のノードの運用管理を管理端末装置で一括して行う。
【解決手段】　複数のノードに、各ノードの主電源により動作し並列処理を実行するメインプロセッサと、各ノードの補助電源により動作し前記メインプロセッサを管理するシステム制御コマンドを実行するサブプロセッサと、前記補助電源により動作し管理端末装置との通信を行うシステム制御機構とを備え、管理端末装置に、前記複数のノードの複数のシステム制御機構と通信を行うシステム制御機構を備え、前記複数のノードの複数のシステム制御機構と前記管理端末装置のシステム制御機構とを接続して成り、前記メインプロセッサを管理するシステム制御コマンドを前記管理端末装置から前記サブプロセッサに送信するシステム制御インタフェースを備えるものである。

図１



1

【特許請求の範囲】

【請求項１】　複数の計算機であるノードを接続した並列計算機システムを管理する管理端末装置を備えた並列計算機システムの管理装置において、

前記複数のノードは、当該ノードの主電源により動作し並列処理を実行するメインプロセッサと、当該ノードの補助電源により動作し前記メインプロセッサを管理するシステム制御コマンドを実行するサブプロセッサと、前記補助電源により動作し前記管理端末装置との通信を行うシステム制御機構とを備え、

前記管理端末装置は、前記複数のノードの複数のシステム制御機構と通信を行うシステム制御機構を備え、

前記複数のノードの複数のシステム制御機構と前記管理端末装置のシステム制御機構とを接続して成り、前記複数のノードの複数のメインプロセッサを管理するシステム制御コマンドを前記管理端末装置から前記複数のノードの複数のサブプロセッサに送信するシステム制御インタフェースを備えることを特徴とする並列計算機システムの管理装置。

【請求項２】　前記複数のノードのサブプロセッサは、当該ノードの主電源を投入または切断する機能を備え、前記管理端末装置は、前記複数のノードのサブプロセッサに、一括または個別に主電源を投入または切断するシステム制御コマンドを送信する手段を備えることを特徴とする請求項１に記載された並列計算機システムの管理装置。

【請求項３】　前記管理端末装置は、前記複数のノードの主電源を個別に投入するシステム制御コマンドを、予め設定された時間間隔で前記複数のノードのサブプロセッサに個別に送信する手段を備えることを特徴とする請求項２に記載された並列計算機システムの管理装置。

【請求項４】　前記管理端末装置は、前記複数のノードの特定のノードのサブプロセッサに特定のシステム制御コマンドを送信し、予め設定された時間内に前記特定のシステム制御コマンドに対する正常な応答が受信されない場合に、前記特定のノードに異常が発生しているとみなす手段を備えることを特徴とする請求項１に記載された並列計算機システムの管理装置。

【請求項５】　前記複数のノードのシステム制御機構は、当該ノードのメインプロセッサまたはサブプロセッサが動作時に出力するメッセージであるノードメッセージを蓄積する手段を備え、前記管理端末装置は、当該ノードのシステム制御機構に蓄積されたノードメッセージを読み取る手段を備えることを特徴とする請求項１に記載された並列計算機システムの管理装置。

【請求項６】　前記複数のノードのサブプロセッサは、当該ノードのメインメモリまたはレジスタの内容を参照及び更新する手段を備え、前記管理端末装置は、前記複数のノードのサブプロセッサに、当該ノードのメインメモリまたはレジスタの内容を参照または更新するシステム制御コマンドを送信する手段を備えることを特徴とする請求項１に記載された並列計算機システムの管理装置。

2

【請求項７】　前記複数のノードのサブプロセッサは、当該ノードのメインプロセッサをリセットする手段を備え、前記管理端末装置は、当該ノードのサブプロセッサに、当該ノードのメインプロセッサをリセットするシステム制御コマンドを送信する手段を備えることを特徴とする請求項１に記載された並列計算機システムの管理装置。

【請求項８】　前記複数のノードのサブプロセッサは、当該ノードのメインメモリの内容を参照及び更新する手段と、当該ノードのメインプロセッサをリセットする手段とを備え、前記管理端末装置は、当該ノードのメインプロセッサが格納しているメインメモリ中のブートストラップデバイス名を参照及び更新するシステム制御コマンドと、当該ノードのメインプロセッサをリセットするシステム制御コマンドとを送信する手段を備えることを特徴とする請求項１に記載された並列計算機システムの管理装置。

【請求項９】　前記管理端末装置を複数備え、前記複数の管理端末装置のうちの一部の管理端末装置の機能を制限する手段を備えることを特徴とする請求項１に記載された並列計算機システムの管理装置。

【請求項１０】　前記管理端末装置は、補助電源で動作し、特定の信号を入力すると前記管理端末装置の主電源を投入する電源投入論理と、前記電源投入論理により主電源が投入されたときに、前記複数のノードのサブプロセッサに、一括または個別に主電源を投入するシステム制御コマンドを送信する手段を備えることを特徴とする請求項１に記載された並列計算機システムの管理装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、並列計算機システムの管理装置に関し、特に、並列計算機システムを構成する複数のノードのメインプロセッサが動作していない場合であっても前記複数のノードの保守及び管理を行う並列計算機システムの管理装置に適用して有効な技術に関するものである。

【０００２】

【従来の技術】従来、複数の計算機で構成される計算機システムの運用管理方法及びその実施装置について、いくつかのものが提案されている。

【０００３】複数のＵＮＩＸマシンのコンソールを１台にまとめたときに発生する運用と監視の負荷増大を防止する複数のＵＮＩＸマシンの集中運用および監視コンソールディスプレイについては、特開平６－２１４７６３号公報に記載されている。

【０００４】その概要は、複数のＵＮＩＸマシンを集中運用及び管理するサーバーであるセンター・コンソール

に、運用目的別にコマンドの宛先を格納した宛先テーブルを作成しておき、前記宛先テーブルに従ってコマンドを実行するものである。

【０００５】複数の計算機から構成される複合計算機システムにおいて、単一のシステムコンソールにより接続する計算機を切り換えて保守及び操作を行った場合の誤操作を防止する複合計算機システムにおけるコンソール切替制御方式については、特開平５－１２０２４７号公報に記載されている。

【０００６】その概要は、複数の計算機内のサービスプロセッサ同士を切替装置に接続し、前記切替装置にシステムコンソールを接続し、前記システムコンソールからは、計算機を識別する識別子を用いて、メッセージ出力対象の計算機を順次切り替えていくことにより、複数の計算機で１台のシステムコンソールを共有する方式であり、システムコンソールにより保守及び操作を行う際に、操作を行おうとしている計算機の識別子と、システムコンソールに接続されている計算機の識別子を比較し、識別子が一致する場合に操作を実行するものである。

【０００７】分散処理システムを構成する各計算機からのメッセージを集中管理するメッセージ集中管理方式については、特開平５－２０２８１号公報に記載されている。

【０００８】その概要は、ネットワークにて接続された複数の計算機内で集中管理ノードを決定し、その集中管理ノードが監視対象ノードの発行する稼働状況メッセージを集中管理する方式である。

【０００９】

【発明が解決しようとする課題】本発明者は、前記従来技術を検討した結果、以下の問題点を見い出した。

【００１０】すなわち、前記従来の複数の計算機から成る計算機システムの管理装置では、管理対象の計算機上で動作しているネットワークソフトウェアの機能を使用してた為、管理対象の計算機が動作していない場合やオペレーティングシステムが動作していない場合及びネットワークソフトウェアが動作していない場合には、運用管理を行えないという問題があった。

【００１１】前記従来の複数のＵＮＩＸマシンの集中運用および監視コンソールディスプレイを使用する方法では、管理対象となる計算機は、オペレーティングシステムのＵＮＩＸが動作していることが前提となる為、オペレーティングシステムが動作していない場合には、コンソールディスプレイから集中運用および監視ができないという問題があった。

【００１２】前記従来の複合計算機システムにおけるコンソール切替制御方式では、システムコンソールと各々のサービスプロセッサとの間に切替装置が存在している為、切替装置なる特別なハードウェアが必要となるという問題があった。



【００１３】前記従来のメッセージ集中管理方式では、複数の計算機から集中管理ノードにメッセージが送られてくる為、前記集中管理ノードがシステムダウンとなったときには、メッセージの集中管理が行えないという問題と、前記メッセージは、ノードが接続されるネットワーク経由で送信されてくる為、各ノードのオペレーティングシステム及びネットワークが起動されていない場合、集中管理ノードから各ノードの状態を管理することが出来ないという問題があった。

【００１４】本発明の目的は、並列処理を実行するメインプロセッサの動作並びに前記メインプロセッサのオペレーティングシステム及びネットワークソフトウェアの動作とは無関係に、並列計算機システムを構成する複数のノードの運用管理を管理端末装置で一括して行うことが可能な技術を提供することにある。

【００１５】本発明の他の目的は、並列計算機システムを構成する複数のノードの電源の投入または切断を管理端末装置で一括または個別に行うことが可能な技術を提供することにある。

【００１６】本発明の他の目的は、並列計算機システムに電力を供給する電源設備の突入電流を低く抑えることが可能な技術を提供することにある。

【００１７】本発明の他の目的は、並列計算機システムを構成する複数のノードが正常に動作中であるかを管理端末装置で監視することが可能な技術を提供することにある。

【００１８】本発明の他の目的は、並列計算機システムを構成する複数のノードのメインプロセッサがノードメッセージを出力した後にその動作を停止した場合であっても、前記ノードメッセージを管理端末装置で一括して管理することが可能な技術を提供することにある。

【００１９】本発明の他の目的は、並列計算機システムを構成する複数のノードの障害発生時のメインメモリ及びレジスタの内容を管理端末装置で一括して管理することが可能な技術を提供することにある。

【００２０】本発明の他の目的は、並列計算機システムを構成する複数のノードのメインプロセッサのリセットを管理端末装置から一括して行うことが可能な技術を提供することにある。

【００２１】本発明の他の目的は、並列計算機システムを構成する複数のノードの特定のブートストラップデバイスに障害が発生した場合に、管理端末装置からの指示により、ブートストラップデバイスを変更して前記複数のノードのメインプロセッサのシステム立ち上げ処理を行うことが可能な技術を提供することにある。

【００２２】本発明の他の目的は、並列計算機システムを複数の管理端末装置で管理した場合に、前記複数の管理端末装置の動作の競合を防止することが可能な技術を提供することにある。

【００２３】本発明の他の目的は、並列計算機システム

の運用管理を遠隔地から行うことが可能な技術を提供することにある。

【００２４】本発明の前記並びにその他の目的と新規な特徴は、本明細書の記述及び添付図面によって明らかになるであろう。

【００２５】

【課題を解決するための手段】本願において開示される発明のうち、代表的なものの概要を簡単に説明すれば、下記のとおりである。

【００２６】（１）複数の計算機であるノードを接続した並列計算機システムを管理する管理端末装置を備えた並列計算機システムの管理装置において、前記複数のノードは、各ノードの主電源により動作し並列処理を実行するメインプロセッサと、各ノードの補助電源により動作し前記メインプロセッサを管理するシステム制御コマンドを実行するサブプロセッサと、前記補助電源により動作し前記管理端末装置との通信を行うシステム制御機構とを備え、前記管理端末装置は、前記複数のノードの複数のシステム制御機構と通信を行うシステム制御機構を備え、前記複数のノードの複数のシステム制御機構と前記管理端末装置のシステム制御機構とを接続して成り、前記複数のノードの複数のメインプロセッサを管理するシステム制御コマンドを前記管理端末装置から前記複数のノードの複数のサブプロセッサに送信するシステム制御インタフェースを備えるものである。

【００２７】前記（１）の並列計算機システムの管理装置では、管理端末装置から発行されたシステム制御コマンドは、システム制御インタフェースを構成する前記管理端末装置及び前記複数のノードのシステム制御機構を介し、前記複数のノードのサブプロセッサに送られ、前記サブプロセッサに送られたシステム制御コマンドを、前記サブプロセッサで実行することにより、前記メインプロセッサの運用管理を行う。

【００２８】従来の並列計算機システムの管理装置では、前記並列計算機システムの通常業務である並列処理を実行するメインプロセッサで動作している汎用のオペレーティングシステムや、そのオペレーティングシステムの管理下で動作するネットワークソフトウェアを使用して、並列計算機システムを構成する複数のノードの運用管理を行っている。

【００２９】この為、前記従来の並列計算機システムの管理装置を使用する場合には、管理対象である並列計算機システムを構成する複数のノードのメインプロセッサが正常に動作し、前記の汎用のオペレーティングシステムやネットワークソフトウェアが実行中であることが前提条件となり、前記管理対象の複数のノードのメインプロセッサが動作していない場合や前記の汎用のオペレーティングシステムやネットワークソフトウェアが動作していない場合、例えば、並列計算機システムの電源が投入されていない運用開始前の状態、または、オペレーティングシステムやネットワークの構成を変更し、正常に動作するかどうか確かめようとしている状態、或いは、障害の発生により動作しなくなった特定のノードの状況を調査する場合などでは、前記従来の並列計算機システムの管理装置を使用することができなかった。



【００３０】そこで、前記（１）の並列計算機システムの管理装置では、メインプロセッサの動作状況とは無関係に動作可能なサブプロセッサとシステム制御機構を、前記（１）の並列計算機システムを構成する全てのノードに備え、各々のノードのシステム制御機構をネットワーク集線装置等の装置を介し、管理端末装置のシステム制御機構に接続している。

【００３１】前記複数のノードの各ノードに備えられたサブプロセッサ及びシステム制御機構は、メインプロセッサが使用する主電源とは別の補助電源により動作し、また、前記システム制御機構は、メインプロセッサで動作するネットワークソフトウェア及びそのネットワークソフトウェアが使用する通信ケーブルとは別のネットワークソフトウェア及び通信ケーブルを使用して管理端末装置と通信を行う。

【００３２】従って、メインプロセッサが動作していなくても、補助電源によりサブプロセッサ及びシステム制御機構が動作していれば、メインプロセッサの制御を前記管理端末装置から行うことが可能である。

【００３３】以上の様に、前記並列計算機システムの管理装置によれば、各ノードの補助電源で動作し、前記メインプロセッサが使用するネットワークソフトウェア及び通信ケーブルとは別のネットワークソフトウェア及び通信ケーブルを使用して管理端末装置と通信を行うシステム制御機構に、前記管理端末装置からシステム制御コマンドを送信し、前記システム制御コマンドを前記補助電源で動作するサブプロセッサで実行することより複数のノードのメインプロセッサの制御を行うので、並列処理を実行するメインプロセッサの動作並びに前記メインプロセッサのオペレーティングシステム及びネットワークソフトウェアの動作とは無関係に、並列計算機システムを構成する複数のノードの運用管理を管理端末装置で一括して行うことが可能である。

【００３４】（２）前記（１）の並列計算機システムの管理装置において、前記複数のノードのサブプロセッサは、当該ノードの主電源を投入または切断する機能を備え、前記管理端末装置は、前記複数のノードのサブプロセッサに、一括または個別に主電源を投入または切断するシステム制御コマンドを送信する手段を備えるものである。

【００３５】前記（２）の並列計算機システムの管理装置では、前記メインプロセッサは主電源により動作するので、前記サブプロセッサにより前記主電源の投入または切断を行うことにより、前記メインプロセッサへの電源の投入を制御することができる。

【００３６】また、前記管理端末装置は、前記主電源を投入または切断するシステム制御コマンドを、送信先を全てのノードまたは特定のノードに指定したパケットとして、前記システム制御インタフェースを介して、前記複数のノードに一括または個別に送信する。

【００３７】前記（２）の並列計算機システムの管理装置では、前記複数のノードのサブプロセッサ及びシステム制御機構は、補助電源により動作しているので、メインプロセッサに主電源が投入されていない場合でも、動作可能である。

【００３８】以上の様に、前記並列計算機システムの管理装置によれば、管理端末装置からの指示により複数のノードの主電源の投入または切断を行うので、並列計算機システムを構成する複数のノードの電源の投入または切断を管理端末装置で一括または個別に行うことが可能である。

【００３９】（３）前記（２）の並列計算機システムの管理装置において、前記管理端末装置は、前記複数のノードの主電源を個別に投入するシステム制御コマンドを、予め設定された時間間隔で、前記複数のノードのサブプロセッサに個別に送信する手段を備えるものである。

【００４０】前記並列計算機システムを構成する複数のノードの主電源を一斉に投入すると、前記主電源に電力を供給する電源設備に過大な突入電流が流れ、前記電源設備に負担をかけるので、前記（３）の並列計算機システムの管理装置では、前記複数のノードの主電源の投入時刻を、各ノードごとにずらし、前記電源設備の突入電源を低く抑える様にする。

【００４１】これは、前記管理端末装置から予め設定された時間間隔で、前記主電源を投入するシステム制御コマンドを、送信先を特定のノードに指定したパケットとして前記システム制御インタフェースを介して送ることにより行われる。

【００４２】以上の様に、前記並列計算機システムの管理装置によれば、前記複数のノードへの主電源の投入指示を、予め設定された特定の時間間隔で行うので、並列計算機システムに電力を供給する電源設備の突入電流を低く抑えることが可能である。

【００４３】（４）前記（１）の並列計算機システムの管理装置において、前記管理端末装置は、前記複数のノードの特定のノードのサブプロセッサに特定のシステム制御コマンドを送信し、予め設定された時間内に前記特定のシステム制御コマンドに対する正常な応答が受信されない場合に、前記特定のノードに異常が発生しているとみなす手段を備えるものである。

【００４４】前記（４）の並列計算機システムの管理装置では、前記特定のシステム制御コマンドを、前記管理端末装置から前記システム制御インタフェースを介して前記サブプロセッサに送信し、前記の送信された特定の



システム制御コマンドを前記サブプロセッサで実行した場合に、前記メインプロセッサの異常により前記特定のシステム制御コマンドの実行結果が得られない場合がある。

【００４５】前記の様な場合に、前記（４）の並列計算機システムの管理装置の管理端末装置は、予め設定された時間の間、前記特定のシステム制御コマンドに対する応答を待ち、前記の予め設定された時間内に前記特定のシステム制御コマンドが正常に実行されたことを示す応答が受信されない場合に、前記特定のノードに異常が発生しているとみなす。

【００４６】以上の様に、前記並列計算機システムの管理装置によれば、管理端末装置からの特定のシステム制御コマンドに対する正常なレスポンスが一定時間中に受信されるかどうかを調べるので、並列計算機システムを構成する複数のノードが正常に動作中であるかを管理端末装置で監視することが可能である。

【００４７】（５）前記（１）の並列計算機システムの管理装置において、前記複数のノードのシステム制御機構は、当該ノードのメインプロセッサまたはサブプロセッサが動作時に出力するメッセージであるノードメッセージを蓄積する手段を備え、前記管理端末装置は、当該ノードのシステム制御機構に蓄積されたノードメッセージを読み取る手段を備えるものである。

【００４８】前記並列計算機システムを構成する複数のノードのメインプロセッサは、各処理の段階で種々のノードメッセージを出力する。

【００４９】例えば、前記並列計算機システムを構成する複数のノードのメインプロセッサは、システム立ち上げ処理中にファイルシステム上に矛盾を発見すると、特定のノードメッセージを出力し、そのファイルシステムの修復を開始する。前記メインプロセッサがファイルシステムの修復に失敗すると、前記ファイルシステムの修復に失敗したことを示すノードメッセージを出力し、前記システム立ち上げ処理は異常終了する。

【００５０】また、前記並列計算機システムを構成する複数のノードのメインプロセッサは、システム立ち上げ処理が正常終了した後、動作中に回復不能な障害を検出すると、パニックメッセージと呼ばれる障害内容や障害発生箇所等の内容を含んだノードメッセージを出力し、前記回復不能な障害を検出したメインプロセッサは、通常、前記パニックメッセージを特定のディスプレイ装置に出力した直後に、システムダウンを起こして動作を停止する。

【００５１】前記の様な場合には、前記ノードメッセージの内容を検討し、システム立ち上げ処理の異常終了やシステムダウンの原因を取り除く必要があるが、当該メインプロセッサは既に動作を停止しているので、従来の並列計算機システムの管理装置では、前記ノードメッセージを管理端末装置に送信して表示することはできなか

った。

【００５２】そこで、前記（５）の並列計算機システムの管理装置では、当該ノードのメインプロセッサまたはサブプロセッサが動作時に出力するノードメッセージを前記複数のノードのシステム制御機構に蓄積し、前記管理端末装置が、当該ノードのシステム制御機構に蓄積されたノードメッセージを読み取ることによって、前記ノードメッセージを前記管理端末装置に表示し、前記ノードメッセージの内容を前記管理端末装置にて検討することを可能にしている。

【００５３】以上の様に、前記並列計算機システムの管理装置によれば、特定のノードのメインプロセッサまたはサブプロセッサが動作時に出力するノードメッセージを蓄積し、前記管理端末装置が前記の蓄積されたノードメッセージを読み取るので、並列計算機システムを構成する複数のノードのメインプロセッサがノードメッセージを出力した後にその動作を停止した場合であっても、前記ノードメッセージを管理端末装置で一括して管理することが可能である。

【００５４】（６）前記（１）の並列計算機システムの管理装置において、前記複数のノードのサブプロセッサは、当該ノードのメインメモリまたはレジスタの内容を参照及び更新する手段を備え、前記管理端末装置は、前記複数のノードのサブプロセッサに、当該ノードのメインメモリまたはレジスタの内容を参照または更新するシステム制御コマンドを送信する手段を備えるものである。

【００５５】前記並列計算機システムを構成する複数のノードのメインプロセッサに障害が発生したときに、当該メインプロセッサに接続されたメインメモリやレジスタの内容を参照して障害の原因を調べ、また、可能な場合には、前記メインメモリやレジスタの内容を変更して前記障害により中断している処理を続行したい場合がある。

【００５６】前記の様な場合、発生した障害によってはメインプロセッサは正常に動作できないことがあるので、前記管理端末装置からのシステム制御コマンドにより、前記障害が発生したメインプロセッサを備えるノードのサブプロセッサが、前記メインメモリまたはレジスタの内容を参照または更新する。

【００５７】以上の様に、前記並列計算機システムの管理装置によれば、管理端末装置からの指示によりノードのメインメモリまたはレジスタの内容を参照または更新するので、並列計算機システムを構成する複数のノードの障害発生時のメインメモリ及びレジスタの内容を管理端末装置で一括して管理することが可能である。

【００５８】（７）前記（１）の並列計算機システムの管理装置において、前記複数のノードのサブプロセッサは、当該ノードのメインプロセッサをリセットする手段を備え、前記管理端末装置は、当該ノードのサブプロセッサに、当該ノードのメインプロセッサをリセットするシステム制御コマンドを送信する手段を備えるものである。

【００５９】前記並列計算機システムを構成する複数のノードにおいて、オペレーティングシステムや他のソフトウェアをバージョンアップしたり、また、障害の原因を取り除く作業を行った後等、メインプロセッサをリセットする必要が生じる場合がある。

【００６０】前記の様な場合に、前記（７）の並列計算機システムの管理装置では、前記管理端末装置からのシステム制御コマンドにより、前記複数のノードのサブプロセッサは、当該ノードのメインプロセッサをリセットする。

【００６１】以上の様に、前記並列計算機システムの管理装置によれば、管理端末装置からの指示により前記複数のノードのメインプロセッサのリセットを行うので、並列計算機システムを構成する複数のノードのメインプロセッサのリセットを管理端末装置から一括して行うことが可能である。

【００６２】（８）前記（１）の並列計算機システムの管理装置において、前記複数のノードのサブプロセッサは、当該ノードのメインメモリの内容を参照及び更新する手段と、当該ノードのメインプロセッサをリセットする手段とを備え、前記管理端末装置は、当該ノードのメインプロセッサが格納しているメインメモリ中のブートストラップデバイス名を参照及び更新するシステム制御コマンドと、当該ノードのメインプロセッサをリセットするシステム制御コマンドとを送信する手段を備えるものである。

【００６３】前記並列計算機システムを構成する複数のノードにおいて、あるメインプロセッサのシステム立ち上げ処理を行うときに、前記システム立ち上げ処理中にメインメモリにロードするオペレーティングシステムや他のソフトウェアを変更する場合や、或いは、オペレーティングシステムや他のソフトウェアを格納しているブートストラップデバイスに障害が発生した場合等、前記システム立ち上げ処理で使用するブートストラップデバイスの変更が必要になることがある。

【００６４】この様な場合に、前記（８）の並列計算機システムの管理装置では、前記管理端末装置により、当該ノードのメインプロセッサが格納しているメインメモリ中のブートストラップデバイス名を参照するシステム制御コマンドを前記サブプロセッサに送り、前記メインメモリ中のブートストラップデバイス名を確認した後、前記管理端末装置は、当該ノードのメインプロセッサが格納しているメインメモリ中のブートストラップデバイス名を他のブートストラップデバイス名に更新するシステム制御コマンドを前記サブプロセッサに送る。

【００６５】前記管理端末装置からメインメモリ中のブートストラップデバイス名を他のブートストラップデバ

イス名に更新するシステム制御コマンドを受け取った前記サブプロセッサは、当該ノードのメインメモリ中のブートストラップデバイス名を更新する。

【００６６】次に、前記管理端末装置は、当該ノードのメインプロセッサをリセットするシステム制御コマンドを前記サブプロセッサに送り、当該ノードのメインプロセッサをリセットして、更新した他のブートストラップデバイスにより前記メインプロセッサのシステム立ち上げ処理を行う。

【００６７】以上の様に、前記並列計算機システムの管理装置によれば、管理端末装置からの指示により前記複数のノードのメインメモリ中のブートストラップパス情報を変更し、メインプロセッサのリセットを行うので、並列計算機システムを構成する複数のノードの特定のブートストラップデバイスに障害が発生した場合に、管理端末装置からの指示により、ブートストラップデバイスを変更して前記複数のノードのメインプロセッサのシステム立ち上げ処理を行うことが可能である。

【００６８】（９）前記（１）の並列計算機システムの管理装置において、前記管理端末装置を複数備え、前記複数の管理端末装置のうちの一部の管理端末装置の機能を制限する手段を備えるものである。

【００６９】前記（９）の並列計算機システムの管理装置では、複数の管理端末装置を備えることにより、特定の管理端末装置が故障した場合に、他の管理端末装置により、前記並列計算機システムの運用管理を行う。

【００７０】前記の様に、前記並列計算機システムに複数の管理端末装置を接続した場合には、前記複数の管理端末装置の動作の内容が、互いに他の管理端末装置の動作の内容と競合する場合がある。

【００７１】この為、前記（９）の並列計算機システムの管理装置では、前記複数の管理端末装置が動作する場合に、特定の管理端末装置をメイン管理端末装置に、他の管理端末装置をサブ管理端末装置に設定し、サブ管理端末装置が行う動作の内容を制限することにより、前記競合の発生を防止する。

【００７２】以上の様に、前記並列計算機システムの管理装置によれば、複数の管理端末装置を備えているので、１つの管理端末装置に障害が発生した場合でも並列計算機システムの運用管理を続行し、並列計算機システムの信頼性を向上させることが可能である。

【００７３】また、前記並列計算機システムの管理装置によれば、複数の管理端末装置にメイン管理端末装置とサブ管理端末装置とを設定するので、並列計算機システムを複数の管理端末装置で管理した場合に、前記複数の管理端末装置の動作の競合を防止することが可能である。

【００７４】（１０）前記（１）の並列計算機システムの管理装置において、前記管理端末装置は、補助電源で動作し、特定の信号を入力すると前記管理端末装置の主電源を投入する電源投入論理と、前記電源投入論理により主電源が投入されたときに、前記複数のノードのサブプロセッサに、一括または個別に主電源を投入するシステム制御コマンドを送信する手段を備えるものである。

【００７５】前記（１０）の並列計算機システムの管理装置では、前記管理端末装置に、補助電源で動作し、特定の信号を入力すると前記管理端末装置の主電源を投入する電源投入論理を接続し、前記電源投入論理をネットワークや他の通信回線に接続しておく。

【００７６】また、前記管理端末装置の主電源が投入されたときに実行されるシステム立ち上げ処理の最後に、前記複数のノードのサブプロセッサに一括または個別に主電源を投入するシステム制御コマンドを送信するプログラムを追加しておく。

【００７７】次に、前記ネットワークや他の通信回線を介して、他の端末装置から前記電源投入論理に特定の信号を送り、前記管理端末装置の主電源を投入する。

【００７８】前記管理端末装置の主電源が投入されると、前記管理端末装置のシステム立ち上げ処理を行った後、前記複数のノードのサブプロセッサに一括または個別に主電源を投入するシステム制御コマンドを送信するプログラムが実行され、前記並列計算機システムの運用開始を、オペレータが直接前記管理端末装置を操作すること無く行うことができる。

【００７９】以上の様に、前記並列計算機システムの管理装置によれば、遠隔地からのアクセスにより管理端末装置の主電源を投入するので、並列計算機システムの運用管理を遠隔地から行うことが可能である。

【００８０】

【発明の実施の形態】以下、本発明について、実施形態とともに図面を参照して詳細に説明する。

【００８１】なお、実施形態を説明するための全図において、同一機能を有するものは同一符号を付け、その繰り返しの説明は省略する。

【００８２】（実施形態１）以下に、本発明の並列計算機システムの管理装置を実施する実施形態１の概略構成について説明する。

【００８３】図１は、本発明の並列計算機システムの管理装置を実施する実施形態１の概略構成を示す図である。図１において、１００ａ～１００ｃはノード、１０１は通信処理装置、１０２ａ～１０２ｃは通信インタフェース機構、１０３ａ～１０３ｃは通信ケーブル、１０４は管理端末装置、１０５ａ～１０５ｄはシステム制御機構、１０６ａ～１０６ｄは通信ケーブル、１０７はネットワーク集線装置、１０８ａ～１０８ｄはＬＡＮ（Ｌｏｃａｌ　Ａｒｅａ　Ｎｅｔｗｏｒｋ）制御機構、１０９ａ～１０９ｄは通信ケーブル、１１０はネットワーク集線装置である。

【００８４】図１に示す様に、本実施形態の並列計算機システムの管理装置は、並列計算機システムを構成する

ノード１００ａ～１００ｃと、並列処理中のノード１００ａ～１００ｃでの通信を制御する通信処理装置１０１と、ノード１００ａ～１００ｃのシステム管理を行う管理端末装置１０４と、管理端末装置１０４とノード１００ａ～１００ｃとを接続するネットワーク集線装置１０７と、ネットワーク集線装置１１０とを備えており、管理端末装置１０４は、システム制御機構１０５ａと、ＬＡＮ制御機構１０８ａとを有し、ノード１００ａは、通信インタフェース機構１０２ａと、システム制御機構１０５ｂと、ＬＡＮ制御機構１０８ｂとを有し、ノード１００ｂは、通信インタフェース機構１０２ｂと、システム制御機構１０５ｃと、ＬＡＮ制御機構１０８ｃとを有し、ノード１００ｃは、通信インタフェース機構１０２ｃと、システム制御機構１０５ｄと、ＬＡＮ制御機構１０８ｄとを有している。

【００８５】また、図１に示す様に、本実施形態の並列計算機システムの管理装置では、ノード１００ａ～１００ｃの通信インタフェース機構１０２ａ～１０２ｃを通信ケーブル１０３ａ～１０３ｃ及び通信処理装置１０１を介して接続し、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄを通信ケーブル１０６ａ～１０６ｄ及びネットワーク集線装置１０７を介して管理端末装置１０４のシステム制御機構１０５ａに接続し、ノード１００ａ～１００ｃのＬＡＮ制御機構１０８ｂ～１０８ｄを通信ケーブル１０９ａ～１０９ｄ及びネットワーク集線装置１１０を介して管理端末装置１０４のＬＡＮ制御機構１０８ａに接続している。

【００８６】本実施形態の並列計算機システムの管理装置のシステム制御インタフェースは、前記の様に、管理端末装置１０４側のシステム制御機構１０５ａとノード１００ａ～１００ｃ側のシステム制御機構１０５ｂ～１０５ｄとをイーサネット等の通信ケーブル１０６及びマルチポートリピータ等のネットワーク集線装置１０７を用いて相互接続することにより実現されるインタフェースである。

【００８７】また、本実施形態の並列計算機システムの管理装置のシステム運用支援インタフェースは、管理端末装置１０４側のＬＡＮ制御機構１０８ａとノード１００ａ～１００ｃ側のＬＡＮ制御機構１０８ｂ～１０８ｄとをイーサネット等の通信ケーブル１０９及びマルチポートリピータ等のネットワーク集線装置１１０を用いて相互接続することにより実現されるインタフェースである。

【００８８】前記システム運用支援インタフェースは、従来の並列計算機システムの運用管理を行うインタフェースであり、ノード１００ａ～１００ｃのメインプロセッサが動作している場合に使用し、ノード１００ａ～１００ｃのメインプロセッサで実行しているアプリケーションソフトウェアが出力するメッセージを管理端末装置１０４に表示する等のシステム管理を行うものである。

【００８９】以下に、本実施形態の並列計算機システムの管理装置において並列計算機システムを構成するノード１００ａ～１００ｃについて説明する。

【００９０】図２は、本実施形態の並列計算機システムの管理装置において並列計算機システムを構成するノード１００ａ～１００ｃの概略構成を示す図である。

【００９１】図２において、２００は主電源、２０１は補助電源、２０２はメインプロセッサ、２０３はソフトウェア、２０４はメインメモリ、２０５はプロセッサメモリ制御機構、２０６はシステムバス、２０７はシステムディスク、２０８はＩ／Ｏ制御機構、２０９はＲＳ－２３２Ｃ制御機構、２１０はブートストラップＲＯＭ（Ｒｅａｄ　Ｏｎｌｙ　Ｍｅｍｏｒｙ）、２１１はシステムサポート機構、２１２はサブプロセッサ、２１３はＲＯＭ、２１４はＳＲＡＭ（Ｓｔａｔｉｃ　Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ；不揮発メモリ）、２１５はローカルバス、２１６は電源投入／切断信号、２１７はプロセッサリセット信号、２１８はＬＡＮ制御部、２１９はＲＳ－２３２Ｃ制御部、２２０はプロセッサ、２２１はＲＯＭ、２２２はＲＡＭ（Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ）、２２３はデータインタフェース、２２４は制御インタフェースである。

【００９２】図２に示す様に、本実施形態の並列計算機システムの管理装置のノード１００ａ～１００ｃは、通信インタフェース機構１０２ａ～１０２ｃと、システム制御機構１０５ｂ～１０５ｄと、ＬＡＮ制御機構１０８ｂ～１０８ｄとを有し、ノード１００ａ～１００ｃで並列処理を行うアプリケーションソフトウェアを実行するメインプロセッサ２０２と、サブプロセッサ２１２を有するシステムサポート機構２１１と、主電源２００と、補助電源２０１とを備えている。

【００９３】また、本実施形態の並列計算機システムの管理装置のノード１００ａ～１００ｃは、メインプロセッサ２０２により実行されるオペレーティングシステム及びネットワークソフトウェアであるソフトウェア２０３と、ソフトウェア２０３を格納するメインメモリ２０４と、メインプロセッサ２０２とメインメモリ２０４とのインタフェース制御を行うプロセッサメモリ制御機構２０５と、システムバス２０６と、システムディスク２０７と、システムディスク２０７を制御するＩ／Ｏ制御機構２０８と、ノードメッセージの出力やシステム制御機構１０５ｂ～１０５ｄ経由のオペレータとのインタラクティブなやりとりを行うＲＳ－２３２Ｃ制御機構２０９と、メインプロセッサ２０２のシステム立ち上げ処理を行うブートストラッププログラムを格納しているブートストラップＲＯＭ２１０とを備えている。

【００９４】本実施形態の並列計算機システムの管理装置において、サブプロセッサ２１２を有し、メインプロセッサ２０２のステータス管理等のシステム制御を行うシステムサポート機構２１１は、サブプロセッサ２１２

15

上で動作する制御プログラムを格納しているＲＯＭ２１３と、ハードウェアに依存した情報を格納しているＳＲＡＭ２１４を備えている。

【００９５】本実施形態の並列計算機システムの管理装置のノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄは、管理端末装置１０４との間でイーサネットパケットの送受信を制御するＬＡＮ制御部２１８と、ＲＳ－２３２Ｃ制御機構２０９及びサブプロセッサ２１２との間でのＲＳ－２３２Ｃパケットの送受信を制御するＲＳ－２３２Ｃ制御部２１９と、イーサネットパケットとＲＳ－２３２Ｃパケットとのプロトコル変換を行うプロセッサ２２０と、プロセッサ２２０上で動作する制御プログラムを格納するＲＯＭ２２１と、サブプロセッサ２１２及びＲＳ－２３２Ｃ制御機構２０９から送られて来るノードメッセージを格納するＲＡＭ２２２とを備えている。

【００９６】図２に示す様に、本実施形態の並列計算機システムの管理装置のノード１００ａ～１００ｃでは、システム制御機構１０５ｂ～１０５ｄを、ＲＳ－２３２Ｃ制御部２１９と、データインタフェース２２３と、ＲＳ－２３２Ｃ制御機構２０９と、システムバス２０６と、プロセッサメモリ制御機構２０５とを介してメインプロセッサ２０２に接続し、また、システム制御機構１０５ｂ～１０５ｄをＲＳ－２３２Ｃ制御部２１９及び制御インタフェース２２４を介してシステムサポート機構２１１のサブプロセッサ２１２に接続し、システムサポート機構２１１のサブプロセッサ２１２を、ローカルバス２１５とプロセッサメモリ制御機構２０５とを介してメインプロセッサ２０２に接続している。また、サブプロセッサ２１２は、プロセッサリセット信号２１７によりメインプロセッサ２０２をリセットし、電源投入／切断信号２１６により主電源２００を制御する。

【００９７】尚、図２に示す様に、本実施形態の並列計算機システムの管理装置のノード１００ａ～１００ｃにおいて、システム制御機構１０５ｂ～１０５ｄを、ＲＳ－２３２Ｃ制御部２１９と、データインタフェース２２３と、ＲＳ－２３２Ｃ制御機構２０９とを介してメインプロセッサ２０２に接続しているのは、システム制御機構１０５ｂ～１０５ｄとメインプロセッサ２０２との間をＲＳ－２３２Ｃ等のシリアルインタフェースで接続することによりその通信ソフトウェアをコンパクトなものとし、メインプロセッサ２０２に障害が発生した場合であっても、システム制御機構１０５ｂ～１０５ｄとメインプロセッサ２０２との間の通信が、できるだけ損なわれることの無い様にする為である。

【００９８】本実施形態の並列計算機システムの管理装置のノード１００ａ～１００ｃは、主電源２００で動作する部位と補助電源２０１で動作する部位より構成されている。

【００９９】主電源２００で動作する部位としては、ノード１００ａ～１００ｃのメインプロセッサ２０２、ソフトウェア２０３を格納するメインメモリ２０４、メインプロセッサ２０２とメインメモリ２０４とのインタフェース制御を行うプロセッサメモリ制御機構２０５、ノード１００ａ～１００ｃのメインプロセッサ２０２のシステム立ち上げ処理を行うブートストラッププログラムを格納しているブートストラップＲＯＭ２１０等があり、これらに、システムバス２０６を介して、通信インタフェース機構１０２ａ～１０２ｃ、ＬＡＮ制御機構１０８ｂ～１０８ｄ等が接続され、また、システムディスク２０７はＩ／Ｏ制御機構２０８経由にて接続される。

【０１００】補助電源２０１で動作する部位としては、ノード１００ａ～１００ｃの主電源２００の制御やメインプロセッサ２０２のステータス管理等のシステム制御を行う部位であるシステムサポート機構２１１と、ノード１００ａ～１００ｃと管理端末装置１０４との通信を制御するシステム制御機構１０５ｂ～１０５ｄがある。

【０１０１】サブプロセッサ２１２は、管理端末装置１０４からの指示により電源投入／切断信号２１６を出力することで、主電源２００の制御を行い、プロセッサリセット信号２１７を出力することで、メインプロセッサ２０２をリセットする機能を持つ。

【０１０２】ノード１００ａ～１００ｃのノードメッセージは、メインプロセッサ２０２が動作し、メインプロセッサ２０２を制御するオペレーティングシステム及びネットワークソフトウェアであるソフトウェア２０３が起動されている状態では、データインタフェース２２３を介してＲＳ－２３２Ｃ制御機構２０９からＲＡＭ２２２に蓄積され、ソフトウェア２０３が起動されていない状態では、サブプロセッサ２１２より、制御インタフェース２２４を介してブートストラップメッセージ等がＲＡＭ２２２に蓄積される。

【０１０３】本実施形態の並列計算機システムの管理装置のシステム制御機構１０５ｂ～１０５ｄのプロセッサ２２０は、前記のパケットのプロトコル変換の他に、以下の処理も行う。

【０１０４】すなわち、管理端末装置１０４からのイーサネットパケットを解釈し、パケットの内容に応じた処理を行い、管理端末装置１０４からの指示によりＲＡＭ２２２に格納しているノードメッセージを管理端末装置１０４に送信する処理を行い、サブプロセッサ２１２は、制御インタフェース２２４を介して送られてきたパケットを解釈し、その内容に応じた制御を行う。

【０１０５】以下に、本実施形態の並列計算機システムの管理装置の管理端末装置１０４の概略構成について説明する。

【０１０６】図３は、本実施形態の並列計算機システムの管理装置の管理端末装置１０４の概略構成を示す図である。図３において、３００はプロセッサ、３０１はソフトウェア、３０２はメインメモリ、３０３はブートス

トラップＲＯＭ、３０４はプロセッサメモリ制御機構、３０５はシステムバス、３０６はＩ／Ｏ制御機構、３０７はシステムディスク、３０８、３０９はＲＳ－２３２Ｃ制御機構、３１０はグラフィックス制御機構、３１１はＬＡＮ制御部、３１２はＲＳ－２３２Ｃ制御部、３１３はプロセッサ、３１４はＲＯＭ、３１５はＲＡＭ、３１６は制御インタフェース、３１７はデータインタフェースである。

【０１０７】図３に示す様に、本実施形態の並列計算機システムの管理装置の管理端末装置１０４は、管理端末装置１０４内の全ての処理を制御／統括するプロセッサ３００と、管理端末装置１０４のオペレーティングシステム及びネットワークソフトウェアであるソフトウェア３０１が格納されているメインメモリ３０２と、管理端末装置１０４のシステム立ち上げ処理を行うブートストラッププログラムを格納しているブートストラップＲＯＭ３０３と、プロセッサ３００、メインメモリ３０２及びブートストラップＲＯＭ３０３のインタフェース制御を行うプロセッサメモリ制御機構３０４とを備えている。

【０１０８】また、本実施形態の並列計算機システムの管理装置の管理端末装置１０４は、システムバス３０５と、システムディスク３０７を制御するＩ／Ｏ制御機構３０６と、システムディスク３０７と、ソフトウェア３０１がノード１００ａ～１００ｃに対し電源制御等のシステム制御コマンドを発行する際に使用するＲＳ－２３２Ｃ制御機構３０８と、ノードメッセージの出力やシステム制御機構１０５ａ経由にてオペレータとのインタラクティブなやりとりを行うＲＳ－２３２Ｃ制御機構３０９と、ディスプレイターミナルやキーボード及びマウスといったマンマシンインタフェースを制御するグラフィックス制御機構３１０と、システム制御機構１０５ａとを備えている。

【０１０９】本実施形態の並列計算機システムの管理装置の管理端末装置１０４のシステム制御機構１０５ａは、ノード１００ａ～１００ｃとの間でイーサネットパケットの送受信を制御するＬＡＮ制御部３１１と、ＲＳ－２３２Ｃ制御機構３０８及び３０９との間でのＲＳ－２３２Ｃパケットの送受信を制御するＲＳ－２３２Ｃ制御部３１２と、イーサネットパケットとＲＳ－２３２Ｃパケットとのプロトコル変換を行うプロセッサ３１３と、プロセッサ３１３で動作する制御プログラムを格納するＲＯＭ３１４と、ノード１００ａ～１００ｃより送られてくるノードメッセージを格納するＲＡＭ３１５とを備えている。

【０１１０】また、図３に示す様に、本実施形態の並列計算機システムの管理装置の管理端末装置１０４では、プロセッサ３００をプロセッサメモリ制御機構３０４を介してメインメモリ３０２、ブートストラップＲＯＭ３０３及びシステムバス３０５に接続し、システムディスク３０７をＩ／Ｏ制御機構３０６を介してシステムバス３０５に接続し、ＬＡＮ制御機構１０８ａと、ＲＳ－２３２Ｃ制御機構３０８及び３０９と、グラフィックス制御機構３１０とをシステムバス３０５に接続している。

【０１１１】更に、図３に示す様に、本実施形態の並列計算機システムの管理装置の管理端末装置１０４では、システム制御機構１０５ａのＲＳ－２３２Ｃ制御部３１２を、制御インタフェース３１６及びデータインタフェース３１７を介してＲＳ－２３２Ｃ制御機構３０８及び３０９に接続している。

【０１１２】本実施形態の並列計算機システムの管理装置において、システム制御インタフェースは、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄと管理端末装置１０４のシステム制御機構１０５ａとをイーサネットケーブル等を用いて、相互接続することにより形成されている。

【０１１３】前記システム制御インタフェースは、管理端末装置１０４側のシステム制御機構１０５ａが動作可能な状態であり、ノード１００ａ～１００ｃの補助電源２０１が投入されており、サブプロセッサ２１２及びシステム制御機構１０５ｂ～１０５ｄが動作可能な状態であれば、ノード１００ａ～１００ｃの主電源２００が投入されておらず、すなわちメインプロセッサ２０２が動作しておらず、メインプロセッサ２０２全体を制御するオペレーティングシステム及びネットワークソフトウェアであるソフトウェア２０３が起動されていなくとも使用可能である。

【０１１４】これに対し、システム運用支援インタフェースは、管理端末装置１０４のＬＡＮ制御機構１０８ａとノード１００ａ～１００ｃのＬＡＮ制御機構１０８ｂ～１０８ｄとをイーサネットケーブル等を用いて、相互接続することにより形成されており、前記システム運用支援インタフェースは、ＴＣＰ／ＩＰ（Ｔｒａｎｓｍｉｓｓｉｏｎ　Ｃｏｎｔｒｏｌ　Ｐｒｏｔｏｃｏｌ／Ｉｎｔｅｒｎｅｔ　Ｐｒｏｔｏｃｏｌ）にて使用するため、管理端末装置１０４及びノード１００ａ～１００ｃのオペレーティングシステム及びそのネットワークソフトウェアであるソフトウェア２０３及びソフトウェア３０１が起動され、ＴＣＰ／ＩＰをサポートするネットワークソフトウェアを実行している状態でのみ使用可能となる。

【０１１５】以下に、本実施形態の並列計算機システムの管理装置の管理端末装置１０４とノード１００ａ～１００ｃとの通信シーケンスについて説明する。

【０１１６】図４は、本実施形態の並列計算機システムの管理装置の管理端末装置１０４とノード１００ａ～１００ｃとの通信シーケンスの一例を示す図である。図４において、４０１はアダプタ制御コマンド及びそのレスポンス、４０２はシステム制御コマンド及びそのレスポンス、４０３はノードメッセージである。

【０１１７】図４に示す様に、本実施形態の並列計算機システムの管理装置では、アダプタ制御コマンド及びそのレスポンス４０１、または、システム制御コマンド及びそのレスポンス４０２であるパケットの送受信、或いは、ノードメッセージ４０３の送受信により通信を行う。

【０１１８】アダプタ制御コマンド及びそのレスポンス４０１は、管理端末装置１０４のソフトウェア３０１が管理端末装置１０４のシステム制御機構１０５ａと通信を行う際、およびサブプロセッサ２１２がシステム制御機構１０５ｂ～１０５ｄと通信を行う際に使用し、制御インタフェース３１６または制御インタフェース２２４を介して送受信される。

【０１１９】システム制御コマンド及びそのレスポンス４０２は、管理端末装置１０４のソフトウェア３０１がノード１００ａ～１００ｃのサブプロセッサ２１２と通信を行う際に使用し、制御インタフェース３１６及び制御インタフェース２２４を介して送受信される。

【０１２０】ノードメッセージ４０３は、ソフトウェア２０３が起動していないときは、サブプロセッサ２１２からシステム制御機構１０５ｂ～１０５ｄのＲＡＭ２２２へ送信されて蓄積され、また、ソフトウェア２０３が起動されているときは、メインプロセッサ２０２からＲＳ－２３２Ｃ制御機構２０９よりシステム制御機構１０５ｂ～１０５ｄのＲＡＭ２２２へ送信されて蓄積される。

【０１２１】システム制御機構１０５ｂ～１０５ｄのＲＡＭ２２２に蓄積されたノードメッセージ４０３は、管理端末装置１０４からの要求により、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄのＲＡＭ２２２から、管理端末装置１０４のシステム制御機構１０５ａを介し、管理端末装置１０４のＲＳ－２３２Ｃ制御機構３０９へ送信され、管理端末装置１０４のグラフィックス制御機構３１０に接続されるグラフィックスディスプレイ等に表示される。

【０１２２】以下に、本実施形態の並列計算機システムの管理装置におけるアダプタ制御コマンド及びそのレスポンス４０１のパケットフォーマットについて説明する。

【０１２３】図５は、本実施形態の並列計算機システムの管理装置におけるアダプタ制御コマンド及びそのレスポンス４０１のパケットフォーマットを示す図である。図５において、５０１は種別フィールド、５０２は送信元アドレスフィールド、５０３は受信先アドレスフィールド、５０４は情報部フィールド、５０５は識別子である。

【０１２４】図５に示す様に、本実施形態の並列計算機システムの管理装置におけるアダプタ制御コマンド及びそのレスポンス４０１のパケットは、種別フィールド５０１と、送信元アドレスフィールド５０２と、受信先アドレスフィールド５０３と、情報部フィールド５０４と、識別子５０５とを備えている。

【０１２５】本実施形態の並列計算機システムの管理装置において、種別フィールド５０１にはアダプタ制御コマンドまたはそのレスポンスであることを示すパケット識別子、例えば「Ａ」が格納され、送信元アドレスフィールド５０２にはパケットの送信元アドレス、受信先アドレスフィールド５０３にはパケットの受信先アドレスが格納される。

【０１２６】また、情報部フィールド５０４には、パケットの種類により、異なったパラメータが格納され、さらにパケットの末尾には、パケットの終わりを示す識別子５０５、例えば「ＬＦ」（ラインフィード）が付加される。

【０１２７】以下に、本実施形態の並列計算機システムの管理装置におけるシステム制御コマンド及びそのレスポンス４０２のパケットフォーマットについて説明する。

【０１２８】図６は、本実施形態の並列計算機システムの管理装置におけるシステム制御コマンド及びそのレスポンス４０２のパケットフォーマットを示す図である。図６において、６０１は種別フィールド、６０２は送信元アドレスフィールド、６０３は受信先アドレスフィールド、６０４は情報部フィールド、６０５は識別子である。

【０１２９】図６に示す様に、本実施形態の並列計算機システムの管理装置におけるシステム制御コマンド及びそのレスポンス４０２のパケットは、種別フィールド６０１と、送信元アドレスフィールド６０２と、受信先アドレスフィールド６０３と、情報部フィールド６０４と、識別子６０５とを備えている。

【０１３０】本実施形態の並列計算機システムの管理装置において、種別フィールド６０１には、システム制御コマンドまたはそのレスポンスであることを示すパケット識別子、例えば「ｄ」が格納され、送信元アドレスフィールド６０２にはパケットの送信元アドレス、受信先アドレスフィールド６０３にはパケットの受信先アドレスが格納される。

【０１３１】また、情報部フィールド６０４には、パケットの種別により異なったパラメータが格納され、さらにパケットの末尾には、パケットの終わりを示す識別子６０５、例えば「ＬＦ」が付加される。

【０１３２】また、本実施形態の並列計算機システムの管理装置において、管理端末装置１０４からの送信パケットの受信先アドレスフィールド６０３に１６進数の「０ｘｆｆｆｆｆｆｆｆ」が格納されると、そのパケットはブロードキャストパケットとなり、全てのノード１００ａ～１００ｃで受信される。

【０１３３】尚、本実施形態の並列計算機システムの管理装置において、「０ｘ」が付加された数字は１６進数

を示すものとする。

【０１３４】以下に、本実施形態の並列計算機システムの管理装置におけるシステム制御機構１０５ａ～１０５ｄの、パケットモードと非パケットモードのモード遷移について説明する。

【０１３５】図７は、本実施形態の並列計算機システムの管理装置におけるシステム制御機構のモード遷移を示す図である。図７において、７０１はパケットモード、７０２は非パケットモード、７０３は「ＳＥＴ－ＭＯＤＥ」コマンドである。

【０１３６】図７に示す様に、本実施形態の並列計算機システムの管理装置におけるシステム制御機構は、固定長のパケットの送受信を行うパケットモード７０１と、不定長のノードメッセージ４０３の送受信を行う非パケットモード７０２とを備え、パケットモード７０１と非パケットモード７０２のモード遷移は、サブプロセッサ２１２からのアダプタ制御コマンドである「ＳＥＴ－ＭＯＤＥ」コマンド７０３を実行することにより行う。

【０１３７】前記の様に、本実施形態の並列計算機システムの管理装置のシステム制御機構１０５ａ～１０５ｄの動作モードは、パケットモード７０１及び非パケットモード７０２の２種類があり、パケットモード７０１は、管理端末装置１０４と複数のノード１００ａ～１００ｃが通信を行う際に設定されるモードであり、非パケットモード７０２は、特定のノードとコネクション型通信を行い、前記特定のノードからのノードメッセージ４０３を管理端末装置１０４に表示する際に設定されるモードである。

【０１３８】尚、本実施形態の並列計算機システムの管理装置において、管理端末装置１０４及びノード１００ａ～１００ｃのシステム制御機構１０５ａ～１０５ｄは、補助電源２０１投入時にはパケットモード７０１にて動作するものとする。

【０１３９】以下に、本実施形態の並列計算機システムの管理装置におけるシステム制御機構１０５ａ～１０５ｄの非パケットモード７０２でのコネクション状態の遷移について説明する。

【０１４０】図８は、本実施形態の並列計算機システムの管理装置におけるシステム制御機構の非パケットモード７０２でのコネクション状態の遷移を示す図である。図８において、８００はディスコネクト状態、８０１はウェイトコネクト状態、８０２はコネクト状態、８０３は「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンド、８０４は管理端末装置１０４上のシステム制御機構１０５ａとノード１００ａ～１００ｃ上のシステム制御機構１０５ｂ～１０５ｃとの間で行われる呼制御である。

【０１４１】図８に示す様に、本実施形態の並列計算機システムの管理装置におけるシステム制御機構の非パケットモード７０２でのコネクション状態には、相手のシステム制御機構が接続されておらずＲＡＭ２２２にノードメッセージ４０３を蓄積していない状態であるディスコネクト状態８００と、相手のシステム制御機構が接続されていないがノードメッセージ４０３をＲＡＭ２２２に蓄積中である状態のウェイトコネクト状態８０１と、相手のシステム制御機構が接続されているコネクト状態８０２とがあり、前記コネクション状態の遷移は、「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンド８０３またはシステム制御機構からの呼制御８０４により行う。

【０１４２】図８に示す様に、本実施形態の並列計算機システムの管理装置において、非パケットモード７０２設定時には、ディスコネクト状態８００、ウェイトコネクト状態８０１及びコネクト状態８０２の３つのコネクト状態を保持し、ディスコネクト状態８００では、システム制御機構同士の通信は不可となり、ウェイトコネクト状態８０１では、相手のシステム制御機構との通信は不可であるが、ノードメッセージ４０３は、ＲＡＭ２２２内に順次蓄積される。

【０１４３】通信を行うシステム制御機構同士がコネクト状態８０２にあるとき、非パケットモード７０２でのコネクション型通信が可能となる。

【０１４４】これらの状態は、「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンド８０３を発行することにより遷移する。また、相手のシステム制御機構からの呼制御８０４によるコネクト要求があった場合、ウェイトコネクト状態８０１からコネクト状態８０２に遷移する。

【０１４５】本実施形態の並列計算機システムの管理装置にて使用するアダプタ制御コマンド及びそのレスポンス４０１の一覧を表１に示す。表１において、情報部は情報部フィールド５０４に格納される情報を示しており、情報部のバイト０の数字は、パケット種別を示す番号である。

【０１４６】

【表１】

表1

| 項番 | パケット種別 | 用途 | 情報部 | | |
|---|---|---|---|---|---|
| | | | ﾊﾞｲﾄ0 | ﾊﾞｲﾄ1 | ﾊﾞｲﾄ2以降 |
| 1 | SET-ADDRESSｺﾏﾝﾄﾞ | ｼｽﾃﾑ制御機構105の初期設定（論理ｱﾄﾞﾚｽの設定等） | 1 | 無し | 無し |
| 2 | SET-ADDRESSﾚｽﾎﾟﾝｽ | 初期設定結果応答 | 1 | 完了ｺｰﾄﾞ | ｽﾃｰﾀｽ情報 |
| 3 | SET-MODEｺﾏﾝﾄﾞ | ｼｽﾃﾑ制御機構105の動作ﾓｰﾄﾞ設定 | 3 | 動作ﾓｰﾄﾞ | 無し |
| 4 | SET-MODEﾚｽﾎﾟﾝｽ | 動作ﾓｰﾄﾞ設定結果の応答 | 3 | 完了ｺｰﾄﾞ | 無し |
| 5 | SET-CONNECTｺﾏﾝﾄﾞ | 非ﾊﾟｹｯﾄﾓｰﾄﾞ702 でのｺﾈｸｼｮﾝ制御 | 5 | ｺﾈｸﾄ指示 | 無し |
| 6 | SET-CONNECTﾚｽﾎﾟﾝｽ | ｺﾈｸｼｮﾝ制御結果 | 5 | 完了ｺｰﾄﾞ | 詳細情報 |
| 7 | REPORT-CONNECT ｲﾝﾃﾞｨｹｰｼｮﾝ | ｺﾈｸﾄ状態変更通知(相手のｼｽﾃﾑ制御機構105からのｺﾈｸﾄ要求を自ﾎﾌﾟﾛｾｯｻ212 に伝える) | A | ｺﾈｸﾄ状態変化状況 | 無し |
| 8 | REPORT-CONNECT ﾚｽﾎﾟﾝｽ | REPORT-CONNECT 受領応答 | A | 無し | 無し |

【０１４７】本実施形態の並列計算機システムの管理装置にて使用するシステム制御コマンド及びそのレスポンス４０２の一覧を表２に示す。表２において、情報部は情報部フィールド６０４に格納される情報を示しており、情報部のバイト０の数字は、パケット種別を示す番号である。

【０１４８】

【表２】

表2

| 項番 | パケット種別 | 用途 | 情報部 | | |
|---|---|---|---|---|---|
| | | | ﾊﾞｲﾄ0 | ﾊﾞｲﾄ1 | ﾊﾞｲﾄ2以降 |
| 1 | P-ONｺﾏﾝﾄﾞ | ノード１００の電源投入 | 1 | 無し | 無し |
| 2 | P-ONﾚｽﾎﾟﾝｽ | 電源投入の結果応答 | 1 | 完了ｺｰﾄﾞ | 無し |
| 3 | P-OFFｺﾏﾝﾄﾞ | ノード１００の電源切断 | 2 | 無し | 無し |
| 4 | P-OFFﾚｽﾎﾟﾝｽ | 電源切断の結果応答 | 2 | 完了ｺｰﾄﾞ | 無し |
| 5 | PROC-RESETｺﾏﾝﾄﾞ | メインプロセッサ２０２リセット | 3 | 無し | 無し |
| 6 | PROC-RESETﾚｽﾎﾟﾝｽ | リセットの結果 | 3 | 完了ｺｰﾄﾞ | 無し |
| 7 | STATUS-READｺﾏﾝﾄﾞ | ノード１００のステータスコードの読み取り | 4 | 無し | 無し |
| 8 | STATUS-READﾚｽﾎﾟﾝｽ | ステータスコードの読み取り結果 | 4 | 完了ｺｰﾄﾞ | ステータス情報 |
| 9 | MS-READｺﾏﾝﾄﾞ | メインメモリ２０４の内容の読み込み | 5 | 先頭アドレス：読み出し長さ | |
| １０ | MS-READﾚｽﾎﾟﾝｽ | 読み出したメモリの値を応答 | 5 | 完了ｺｰﾄﾞ | 読み出した値 |
| １１ | MS-WRITEｺﾏﾝﾄﾞ | メインメモリ２０４への書き込み | 6 | 先頭アドレス：書き込みデータ | |
| １２ | MS-WRITEﾚｽﾎﾟﾝｽ | メインメモリ２０４へ書き込み後の値を応答 | 6 | 完了ｺｰﾄﾞ | 書き込み後の値 |
| １３ | REG-READｺﾏﾝﾄﾞ | レジスタの内容の読み込み | 7 | レジスタアドレス：読み出し長さ | |
| １４ | REG-READﾚｽﾎﾟﾝｽ | 読み出したレジスタの値を応答 | 7 | 完了ｺｰﾄﾞ | 読み出した値 |
| １５ | REG-WRITEｺﾏﾝﾄﾞ | レジスタへの書き込み | 8 | レジスタアドレス：書き込みデータ | |
| １６ | REG-WRITEﾚｽﾎﾟﾝｽ | レジスタへ書き込み後の値を応答 | 8 | 完了ｺｰﾄﾞ | 書き込み後の値 |

【０１４９】以下に、本実施形態の並列計算機システム　50　の管理装置におけるノード１００ａ～１００ｃのシステ

ム制御機構１０５ｂ～１０５ｄのプロセッサ２２０の処理手順について説明する。

【０１５０】図９は、本実施形態の並列計算機システムの管理装置におけるノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄのプロセッサ２２０の処理手順の一部を示すフローチャートである。

【０１５１】図９に示す様に、本実施形態の並列計算機システムの管理装置におけるノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄのプロセッサ２２０では、ステップ９００の処理にて、「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンドや呼制御により管理端末装置１０４からコネクト要求があるかどうかを調べる。

【０１５２】ステップ９００の処理で、「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンドや呼制御により管理端末装置１０４からのコネクト要求がある場合には、ステップ９０１の処理に進み、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄが非パケットモード７０２であるかどうかをチェックする。

【０１５３】ステップ９０１の処理で、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄが非パケットモード７０２であれば、ステップ９０２の処理へ進み、ステップ９０１の処理で、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄが非パケットモード７０２でなければ、ステップ９０３の処理にて、サブプロセッサ２１２からのシステム制御コマンド「ＳＥＴ－ＭＯＤＥ」により、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄを非パケットモード７０２に設定し、ステップ９０２の処理へ進む。

【０１５４】ステップ９０２の処理では、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄのＲＡＭ２２２に蓄積されたノードメッセージ４０３をシステム制御インタフェース経由で管理端末装置１０４へ送信し、ステップ９００の処理に戻る。

【０１５５】ステップ９００の処理にて「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンドや呼制御により管理端末装置１０４からコネクト要求が無い場合には、ステップ９０４の処理に進み、ステップ９０４の処理にて、システム制御コマンドにより、管理端末装置１０４からのシステム制御があるかどうかを調べる。

【０１５６】ステップ９０４の処理にて、前記システム制御コマンドにより、管理端末装置１０４からのシステム制御がある場合には、ステップ９０５の処理に進み、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄがパケットモード７０１かどうかをチェックする。

【０１５７】ステップ９０４の処理にて、システム制御コマンドによる管理端末装置１０４からのシステム制御がない場合には、ステップ９０９の処理に進む。

【０１５８】ステップ９０５の処理で、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄがパケットモード７０１であれば、ステップ９０６の処理へ進み、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄがパケットモード７０１でなければ、ステップ９０７の処理にて、サブプロセッサ２１２からのシステム制御コマンド「ＳＥＴ－ＭＯＤＥ」により、システム制御機構１０５ｂ～１０５ｄをパケットモード７０１に設定し、ステップ９０６の処理へ進む。

【０１５９】ステップ９０６の処理にて、前記システム制御コマンドの受信先アドレスフィールド６０３をチェックし、前記システム制御コマンドの受信先アドレスフィールド６０３が、自論理アドレスまたは「０ｘｆｆｆｆｆｆｆｆ」である場合は、ステップ９０８の処理に進み、前記システム制御コマンドの内容をサブプロセッサ２１２に通知し、ステップ９００の処理に戻る。

【０１６０】ステップ９０６の処理にて、前記システム制御コマンドの受信先アドレスフィールド６０３が、自論理アドレス及び「０ｘｆｆｆｆｆｆｆｆ」でない場合は、ステップ９００の処理に戻る。

【０１６１】ステップ９０９の処理にて、ノード１００ａ～１００ｃのサブプロセッサ２１２からの処理の結果が返ってきたかどうかを調べ、サブプロセッサ２１２からの処理の結果が返ってきた場合には、ステップ９１０の処理に進み、管理端末装置１０４に対し、前記システム制御コマンドのレスポンスパケットを送信し、ステップ９００の処理に戻る。

【０１６２】以下に、本実施形態の並列計算機システムの管理装置におけるノード１００ａ～１００ｃのシステムサポート機構２１１のサブプロセッサ２１２の処理手順について説明する。

【０１６３】図１０は、本実施形態の並列計算機システムの管理装置におけるノード１００ａ～１００ｃのシステムサポート機構２１１のサブプロセッサ２１２の処理手順の一部を示すフローチャートである。

【０１６４】図１０に示す様に、本実施形態の並列計算機システムの管理装置におけるノード１００ａ～１００ｃのシステムサポート機構２１１のサブプロセッサ２１２では、補助電源２０１が投入されると、ステップ１０００の処理にて、ノード１００ａ～１００ｃの論理アドレスを設定し、ノード１００ａ～１００ｃに備えられたパネルに表示するステータスコードを格納するＳＲＡＭ２１４内のパネルステータス管理領域に「００００」を設定する。

【０１６５】次に、ステップ１００１の処理にて、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄを非パケットモード７０２に設定し、ステップ１００２の処理にて、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄの非パケットモード７０２のコネクション状態をウェイトコネクト状態８０１に設定する。

【０１６６】ノード１００ａ～１００ｃのシステム制御

機構１０５ｂ～１０５ｄのモードを非パケットモード７０２に設定し、システム制御機構１０５ｂ～１０５ｄの非パケットモード７０２のコネクション状態をウェイトコネクト状態８０１に設定するのは、ノード１００ａ～１００ｃのノードメッセージをシステム制御機構１０５ｂ～１０５ｄのＲＡＭ２２２に蓄積すると共に、管理端末装置１０４のシステム制御機構１０５ａからの呼制御８０４によるコネクト要求があったときに、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄのＲＡＭ２２２に蓄積したノードメッセージを管理端末装置１０４に送る為である。

【０１６７】また、こうすることでノード１００ａ～１００ｃ上のソフトウェア２０３が起動されていない場合でも管理端末装置１０４からＲＡＭ２２２に蓄積したノードメッセージを読み出すことが可能となる。

【０１６８】次に、管理端末装置１０４のシステム制御機構１０５ａから、ノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄにシステム制御コマンドが送られた場合には、前記システム制御コマンドをノード１００ａ～１００ｃのサブプロセッサ２１２に送り、サブプロセッサ２１２にて前記システム制御コマンドを実行する。

【０１６９】ステップ１００３の処理にて、管理端末装置１０４のシステム制御機構１０５ａからノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄを介して、ノード１００ａ～１００ｃの主電源２００を投入または切断する電源制御指示のシステム制御コマンドが送られてきているかどうかを調べる。

【０１７０】ステップ１００３の処理で管理端末装置１０４からの電源制御指示があるかどうかを調べた結果、管理端末装置１０４からの電源制御指示がある場合には、ステップ１００４の処理にて、ノード１００ａ～１００ｃの主電源２００を投入または切断する電源制御処理を実行し、ステップ１００５の処理にて、前記電源制御処理の実行結果をノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄへ報告した後、ステップ１００３の処理に戻る。

【０１７１】ステップ１００３の処理で管理端末装置１０４からの電源制御指示があるかどうかを調べた結果、管理端末装置１０４からの電源制御指示がない場合には、ステップ１００６の処理に進み、管理端末装置１０４のシステム制御機構１０５ａからノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄを介して、ノード１００ａ～１００ｃに備えられたパネルを制御するパネル制御指示のシステム制御コマンドが送られてきているかどうかを調べる。

【０１７２】ステップ１００６の処理にて、管理端末装置１０４からのパネル制御指示があるかどうかを調べた結果、管理端末装置１０４からのパネル制御指示がある場合には、ステップ１００７の処理に進み、パネル制御処理を実行し、ステップ１００８の処理にて、前記パネル制御処理の実行結果をノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄへ報告した後、ステップ１００３の処理に戻る。

【０１７３】ステップ１００６の処理にて、管理端末装置１０４からのパネル制御指示があるかどうかを調べた結果、管理端末装置１０４からのパネル制御指示がない場合には、ステップ１００９の処理に進み、管理端末装置１０４のシステム制御機構１０５ａからノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄを介して、ノード１００ａ～１００ｃのメインプロセッサ２０２をリセットするリセット指示のシステム制御コマンドが送られてきているかどうかを調べる。

【０１７４】ステップ１００９の処理にて、管理端末装置１０４からのリセット指示があるかどうかを調べた結果、管理端末装置１０４からのリセット指示がある場合には、ステップ１０１０の処理に進み、ノード１００ａ～１００ｃのメインプロセッサ２０２のリセット処理を実行し、ステップ１０１１の処理にて、前記リセット処理の実行結果をノード１００ａ～１００ｃのシステム制御機構１０５ｂ～１０５ｄへ報告した後、ステップ１００３の処理に戻る。

【０１７５】ステップ１０１２の処理にてシステム制御機構１０５ｂ～１０５ｄからモード切り替えの要求があるかどうかを調べた結果、モード切り替え要求がある場合には、ステップ１０１３の処理に進み、アダプタ制御コマンド「ＳＥＴ－ＭＯＤＥ」を実行し、システム制御機構１０５ｂ～１０５ｄの動作モードを切り替え、ステップ１００３の処理に戻る。

【０１７６】以上説明した様に、本実施形態の並列計算機システムの管理装置によれば、ノード１００ａ～１００ｃの補助電源２０１で動作し、メインプロセッサ２０２が使用するネットワークソフトウェア及び通信ケーブル１０９ｂ～１０９ｄとは別のネットワークソフトウェア及び通信ケーブル１０６ｂ～１０６ｄを使用して管理端末装置１０４と通信を行うシステム制御機構１０５ｂ～１０５ｄに、管理端末装置１０４からシステム制御コマンドを送信し、前記システム制御コマンドを補助電源２０１で動作するサブプロセッサ２１２で実行することより、複数のノード１００ａ～１００ｃのメインプロセッサ２０２の制御を行うので、並列処理を実行するメインプロセッサ２０２の動作並びにメインプロセッサ２０２のオペレーティングシステム及びネットワークソフトウェアであるソフトウェア２０３の動作とは無関係に、並列計算機システムを構成する複数のノード１００ａ～１００ｃの運用管理を管理端末装置１０４で一括して行うことが可能である。

【０１７７】（実施形態２）以下に、本発明の並列計算機システムの管理装置において、管理端末装置１０４から複数のノード１００ａ～１００ｄに主電源２００の投

入を指示し、ノード１００ａ～１００ｄのステータスコードを監視し、ノード１００ａ～１００ｄのメインプロセッサ２０２が動作を開始したかどうかを管理する実施形態２について説明する。

【０１７８】図１１は、本実施形態の並列計算機システムの管理装置における管理端末装置１０４からノード１００ａ～１００ｄへ主電源２００の投入を指示する電源投入シーケンスの一例を示す図である。図１１において、１００ｄはノード、１１０１～１１１２は電源投入の各段階を示すシーケンスである。

【０１７９】図１１に示す様に、本実施形態の並列計算機システムの管理装置における管理端末装置１０４からノード１００ａ～１００ｄへ主電源２００の投入を指示する電源投入シーケンスでは、シーケンス１１０１にて、ノード１００ａ～１００ｄの補助電源２０１が投入されている。

【０１８０】ノード１００ａ～１００ｄの補助電源２０１が投入されると、ノード１００ａ～１００ｄのサブプロセッサ２１２は、シーケンス１１０２にて、システムサポート機構２１１内の初期化を行い、アダプタ制御コマンド「ＳＥＴ－ＡＤＤＲＥＳＳ」によって、システム制御機構１０５ｂ～１０５ｄの初期化、及び、管理端末装置１０４がノード１００ａ～１００ｄを管理するために必要なアドレスである論理アドレスの設定を行う。

【０１８１】ここで、例えば、論理アドレス「０ｘ０００００００１」を設定する「ＳＥＴ－ＡＤＤＲＥＳＳ」コマンド及びそのレスポンスのフォーマットの一例は、下記の通りとなる。

【０１８２】＜コマンド＞：
A0x00000001：（受信先アドレスフィールド５０３は省略）：0x01 LF
＜レスポンス＞：
A0x00000001：0x00000001：0x01（ステータス情報）LF
シーケンス１１０３にて、管理端末装置１０４の電源が投入されると、管理端末装置１０４のブートストラップＲＯＭ３０３に格納されているブートストラッププログラムが、管理端末装置１０４のシステム立ち上げ処理を行う。

【０１８３】シーケンス１１０４にて、管理端末装置１０４のシステム立ち上げ処理が終わると、シーケンス１１０５にて、管理端末装置１０４のソフトウェア３０１は、管理端末装置１０４の論理アドレスを「ＳＥＴ－ＡＤＤＲＥＳＳ」にて設定する。

【０１８４】管理端末装置１０４及びノード１００ａ～１００ｄの論理アドレスが設定されると、シーケンス１１０６にて、管理端末装置１０４のソフトウェア３０１は、システム制御コマンドのブロードキャストパケットを用いて、ノード１００ａ～１００ｄの状態を示すステータスコードを読み出す。

【０１８５】ステータスコードは、ノード１００ａ～１００ｄのＳＲＡＭ２１４内のパネルステータス管理領域にて管理されており、例えば、ノード１００ａ～１００ｄの補助電源２０１が正常に投入されると、ある一定のステータスコードが前記パネルステータス管理領域に書き込まれ、また、そのステータスコードは、サブプロセッサ２１２により読み出すことができる（本実施形態の並列計算機システムの管理装置ではコード「００００」が読み出せるものとする。）。

【０１８６】ここでは、管理端末装置１０４は「ＳＴＡＴＵＳ－ＲＥＡＤ」コマンドを使用して、ノード１００ａ～１００ｄに対し、ブロードキャストを行う。

【０１８７】論理アドレスが「０ｘａ０００００００」である管理端末装置１０４が、「ＳＴＡＴＵＳ－ＲＥＡＤ」コマンドをブロードキャストした場合と、そのコマンドに対する、論理アドレスが「０ｘ００００００００５」であるノードからのレスポンスのフォーマットの一例は、下記の通りとなる。

【０１８８】＜コマンド＞：
d 0xa0000000：0xffffffff：0x4 LF
＜レスポンス＞
d 0x00000005：0xa0000000：0x04 0000 LF
シーケンス１１０７にて、ノード１００ａ～１００ｄで前記「ＳＴＡＴＵＳ－ＲＥＡＤ」コマンドが受信され、サブプロセッサ２１２によりステータスコード「００００」が読み出された後、シーケンス１１０８にて、ノード１００ａ～１００ｄから管理端末装置１０４に対し、前記の様にレスポンスが返ってくる。

【０１８９】ここで、管理端末装置１０４のソフトウェア３０１は、正常なレスポンスが返ってきたノードの論理アドレスと、予め管理端末装置１０４のソフトウェア３０１内または特定のファイルに保持しておいた、並列計算機システムを構成するノード１００ａ～１００ｄの構成情報とを照らし合わせ、正常なレスポンスが返ってこないノードに対しては、予め設定された一定の時間間隔で再び「ＳＴＡＴＵＳ－ＲＥＡＤ」コマンドを送るリトライ処理を行う。

【０１９０】シーケンス１１０９にて、管理端末装置１０４のソフトウェア３０１は、シーケンス１１０８で正常なレスポンスパケットが返ってきたノードの主電源２００を「Ｐ－ＯＮ」コマンドにて投入する。

【０１９１】例えば、論理アドレスが「０ｘａ００００００００」である管理端末装置１０４から、論理アドレスが「０ｘ００００００００５」であるノードに対する「Ｐ－ＯＮ」コマンド及びそのレスポンスのフォーマットの一例は、下記の通りとなる。

【０１９２】＜コマンド＞：
d 0xa0000000：0x00000005：0x01LF
＜レスポンス＞：
d 0x00000005：0xa0000000：0x01（完了コード）LF
このとき、管理端末装置１０４のソフトウェア３０１の

制御により、予め設定された一定の時間間隔で「Ｐ－ＯＮ」コマンドをずらしながらノード１００ａ～１００ｄに送信することで、並列計算機システム全体に電源を供給している電源設備への突入電流を低く抑えることが出来る。

【０１９３】シーケンス１１１０にて、「Ｐ－ＯＮ」コマンドを受け取ったノード１００ａ～１００ｄのサブプロセッサ２１２は、電源投入信号２１６を出力し、主電源２００をオンにした後、「Ｐ－ＯＮ」コマンドに対するレスポンスを、管理端末装置１０４に返送する。

【０１９４】ノード１００ａ～１００ｄの主電源２００＊がオンになると、メインプロセッサ２０２によりブートストラップＲＯＭ２１０に格納されているブートストラッププログラムが実行され、システム立ち上げ処理が開始される。

【０１９５】尚、システム立ち上げ処理中にブートストラッププログラムがインクリメントするステータスコードには、例えば以下のようなものがある。ここで、本実施形態の並列計算機システムの管理装置では、ステータスコードは１６進数で示されている。

【０１９６】

【表３】

表３

| ステータスコード | 意　味 |
|---|---|
| １０００ | ハードウェアの初期化中 |
| １ＦＦＦ | ハードウェア障害 |
| ２０００ | プログラムのロード中 |
| ２ＦＦＦ | プログラムロードエラー |
| ３０００ | ハードウェアブートストラップ終了 |
| Ａ０００ | ソフトウェアの初期化中 |
| Ｆ０００ | アプリケーションソフトウェア動作中 |

【０１９７】ノード１００ａ～１００ｄのメインプロセッサ２０２のブートストラッププログラムは、ノード１００ａ～１００ｄのＳＲＡＭ２１４内のパネルステータス管理領域にステータスコードを書き込み、システム立ち上げ処理が進むと、定期的に前記ステータスコードを更新する。

【０１９８】また、前記パネルステータス管理領域は、ノード１００ａ～１００ｄのサブプロセッサ２１２からも参照可能であり、例えば、ノード１００ａ～１００ｄに備えられたパネル等の表示装置に表示することにより、オペレータに対し、前記ステータスコードを開示することも可能である。

【０１９９】管理端末装置１０４のソフトウェア３０１は、これらのノード１００ａ～１００ｄのステータスコードを「ＳＴＡＴＵＳ－ＲＥＡＤ」コマンドを使用して定期的に読み出すことにより、ノード１００ａ～１００ｄの状態を監視する。

【０２００】シーケンス１１１０にて、管理端末装置１０４のソフトウェア３０１は、システム制御コマンドの送信からそのレスポンスの受信までを一定の時間で監視しており、図１１に示す様に、何らかの障害が発生しており、一定時間内に正常なレスポンスが返ってこないノード１００ｄに対しては、シーケンス１１１１にて、予め設定された一定の時間間隔で再度システム制御コマンドを送信するリトライ処理を行う。

【０２０１】図１１に示す様に、本実施形態の並列計算機システムの管理装置において、一定回数（本実施形態では3回）のリトライ処理の結果、ノード１００ｄから正常なレスポンスが返って来なかった場合、シーケンス１１１２にて、管理端末装置１０４のソフトウェア３０１は、ノード１００ｄに障害が発生していることを認識する。

【０２０２】管理端末装置１０４のソフトウェア３０１は、前記の様に、特定のシステム制御コマンドに対する正常なレスポンスが一定時間内の間に受信されない場合に、予め設定された一定の時間間隔で前記特定のシステム制御コマンドを再度送信する制御を行うことで、ノード１００ａ～１００ｄのソフトウェア２０３が起動されていなくとも、ノード１００ａ～１００ｄのメインプロセッサ２０２のシステム立ち上げ処理が正常に終了しているかどうかの管理を行うことが可能である。

【０２０３】以上説明した様に、本実施形態の並列計算機システムの管理装置によれば、管理端末装置１０４からの指示により複数のノード１００ａ～１００ｄの主電源２００の投入または切断を行うので、並列計算機システムを構成する複数のノード１００ａ～１００ｄの主電源２００の投入または切断を管理端末装置１０４で一括または個別に行うことが可能である。

【０２０４】また、本実施形態の並列計算機システムの管理装置によれば、ノード１００ａ～１００ｄへの主電源２００の投入指示を、予め設定された特定の時間間隔で行うので、並列計算機システムに電力を供給する電源設備の突入電流を低く抑えることが可能である。

【０２０５】また、本実施形態の並列計算機システムの管理装置によれば、管理端末装置１０４からの指示によりノード１００ａ～１００ｄのステータスコードを読み

出すので、複数のノード１００ａ～１００ｄの状態を管理端末装置１０４で一括して管理することが可能である。

【０２０６】また、本実施形態の並列計算機システムの管理装置によれば、管理端末装置１０４からの特定のシステム制御コマンドに対する正常なレスポンスが一定時間中に受信されるかどうかを調べるので、並列計算機システムを構成する複数のノードが正常に動作中であるかを管理端末装置１０４で監視することが可能である。

【０２０７】（実施形態３）以下に、本発明の並列計算機システムの管理装置において、管理端末装置１０４にノード１００ａからのノードメッセージ４０３を表示し、必要に応じて保守を行う実施形態３について説明する。

【０２０８】図１２は、本実施形態の並列計算機システムの管理装置における管理端末装置１０４にノード１００ａからのノードメッセージ４０３を表示するシーケンスの一例を示す図である。図１２において、１２０１～１２１７はノードメッセージ４０３を表示する各段階のシーケンスを示している。

【０２０９】図１２に示す様に、本実施形態の並列計算機システムの管理装置における管理端末装置１０４にノード１００ａからのノードメッセージ４０３を表示するシーケンスにおいて、シーケンス１２０１では、ノード１００ａには、予め補助電源２０１が投入されており、システム制御機構１０５ｂ（動作モードはパケットモード７０１）、サブプロセッサ２１２及びプロセッサ２２０は動作可能な状態にある。

【０２１０】補助電源２０１が投入されているノード１００ａのサブプロセッサ２１２は、シーケンス１２０２で「ＳＥＴ－ＡＤＤＲＥＳＳ」コマンドにて、ノード１００ａの論理アドレスを設定する。

【０２１１】次に、ノード１００ａのサブプロセッサ２１２は、シーケンス１２０３で、「ＳＥＴ－ＭＯＤＥ」コマンドにてシステム制御機構１０５ｂの動作モードを非パケットモード７０２（ディスコネクト状態８００）に設定する。

【０２１２】ノード１００ａのサブプロセッサ２１２は、シーケンス１２０４で、さらに「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンドにて、コネクション状態を非パケットモード７０２のウェイトコネクト状態８０１に設定する。

【０２１３】一方、管理端末装置１０４は、シーケンス１２０５で、管理端末装置１０４の電源が投入されると、管理端末装置１０４のシステム立ち上げ処理を開始する。

【０２１４】管理端末装置１０４のシステム立ち上げ処理が終了すると、シーケンス１２０６で、管理端末装置１０４のソフトウェア３０１は、ノード１００ａと同様にして、「ＳＥＴ－ＡＤＤＲＥＳＳ」コマンドを用いて管理端末装置１０４の論理アドレスの設定を行い、シーケンス１２０７で、「ＳＥＴ－ＭＯＤＥ」を用いて、動作モードを非パケットモード７０２のディスコネクト状態８００に設定する。

【０２１５】シーケンス１２０８にて、管理端末装置１０４のソフトウェア３０１は、「ＳＴＡＴＵＳ－ＲＥＡＤ」コマンドによってノード１００ａのステータスコードを読み出し、ステータスコード「００００」が読み出せると、シーケンス１２０９にて、「Ｐ－ＯＮ」コマンドをノード１００ａに送信し、ノード１００ａの主電源２００の投入を指示する。

【０２１６】管理端末装置１０４からの「Ｐ－ＯＮ」コマンドを受信し、主電源２００を投入したノード１００ａは、ブートストラップＲＯＭ２１０に格納されているブートストラッププログラムをメインプロセッサ２０２により実行し、ノード１００ａのシステム立ち上げ処理を行う。

【０２１７】このとき、ノード１００ａのブートストラッププログラムから出力されるノードメッセージ４０３は、サブプロセッサ２１２を経由し、ノード１００ａのシステム制御機構１０５ｂのＲＡＭ２２２に蓄積される。

【０２１８】管理端末装置１０４のソフトウェア３０１は、シーケンス１２１０で、「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンドにより、管理端末装置１０４のシステム制御機構１０５ａのコネクション状態をコネクト状態８０２にすることで、ノード１００ａのシステム制御機構１０５ｂのＲＡＭ２２２に蓄積されているノード１００ａのメインプロセッサ２０２のシステム立ち上げ処理中のノードメッセージ４０３の監視を開始する。

【０２１９】「ＳＥＴ－ＣＯＮＮＥＣＴ」を受けた管理端末装置１０４のシステム制御機構１０５ａは、シーケンス１２１１で、ノード１００ａのシステム制御機構１０５ｂと呼制御８０４を行い、これを受けたノード１００ａのシステム制御機構１０５ｂのコネクション状態は、ウェイトコネクト状態８０１からコネクト状態８０２に遷移する。

【０２２０】同時にノード１００ａのシステム制御機構１０５ｂは、シーケンス１２１２で、「ＲＥＰＯＲＴ－ＣＯＮＮＥＣＴ」コマンドを、ノード１００ａのサブプロセッサ２１２に発行し、管理端末装置１０４からのコネクト要求があったことを伝える。

【０２２１】このときの「ＲＥＰＯＲＴ－ＣＯＮＮＥＣＴ」コマンド及びそのレスポンスのフォーマットの一例は、下記の通りとなる。尚、以下の「ＲＥＰＯＲＴ－ＣＯＮＮＥＣＴ」コマンド及びそのレスポンスでは、送受信アドレスは省略されている。

【０２２２】＜コマンド＞：

Ａ：：0xA（コネクト状態変化状況）LF

＜レスポンス＞：

Ａ：：0xA LF
シーケンス１２１３にて、ノード１００ａのノードメッセージ４０３は、ノード１００ａのシステム制御機構１０５ｂが呼制御８０４によるコネクト要求を受け取った時点で、ノード１００ａのシステム制御機構１０５ｂのＲＡＭ２２２からＬＡＮ制御部２１８を経由して管理端末装置１０４へ送信される。

【０２２３】ここで、図１２には特に示していないが、もし、ノード１００ａのメインプロセッサ２０２のシステム立ち上げ処理中に、ノード１００ａのメインプロセッサ２０２が使用するファイルシステムに矛盾が発見され、前記システム立ち上げ処理が中断した場合には、シーケンス１２１４にて、オペレータは、管理端末装置１０４の表示装置に出力されるノード１００ａのメインプロセッサ２０２のシステム立ち上げ処理中のノードメッセージ４０３により、ノード１００ａに障害が発生していることを認識し、ＵＮＩＸのｆｓｃｋ等のファイルシステムを検査する保守コマンドを投入することで、ノード１００ａの保守を行うことも可能である。

【０２２４】また、ノード１００ａのメインプロセッサ２０２のシステム立ち上げ処理は正常終了したが、その後の通常の業務でノード１００ａのメインプロセッサ２０２を使用中に、パニックメッセージを出力してノード１００ａのメインプロセッサ２０２がシステムダウンを起こした場合には、オペレータは、ノード１００ａのシステム制御機構１０５ｂのＲＡＭ２２２に蓄積されたノードメッセージ４０３を管理端末装置１０４に表示させ、ノードメッセージ４０３の内容によりシステムダウンの要因を検討することも可能である。

【０２２５】ノード１００ａとのコネクションを切断する場合には、シーケンス１２１０～１２１２までの処理と同様、管理端末装置１０４のソフトウェア３０１が、シーケンス１２１５にて、「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンドを発行する。

【０２２６】管理端末装置１０４のソフトウェア３０１が「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンドを発行することにより、管理端末装置１０４のシステム制御機構１０５ａはディスコネクト状態８００になり、シーケンス１２１６にて、管理端末装置１０４のシステム制御機構１０５ａは、ノード１００ａのシステム制御機構１０５ｂに対し呼制御８０４を行う。

【０２２７】シーケンス１２１７にて、ノード１００ａのシステム制御機構１０５ｂは、前記の管理端末装置１０４のシステム制御機構１０５ａからの呼制御８０４により、相手からコネクション断の要求があったことを認識し、同時にノード１００ａのサブプロセッサ２１２に対し、このことを「ＲＥＰＯＲＴ－ＣＯＮＮＥＣＴ」コマンドにて報告する。

【０２２８】以上の様に、管理端末装置１０４からノード１００ａ～１００ｃに対し、システム制御コマンドを送信することで、管理端末装置１０４からノード１００ａ～１００ｃの主電源２００の電源制御及びノード１００ａ～１００ｃの状態監視が可能となる。

【０２２９】以上説明した様に、本実施形態の並列計算機システムの管理装置によれば、パケットモード７０１及び非パケットモード７０２のモード切り替えをサブプロセッサ２１２により行うので、パケットを用い、システム制御コマンド及びそのレスポンスを複数のノード１００ａ～１００ｃと送受信する通信と、特定のノードとのコネクションを設定し、特定のノードのノードメッセージ４０３を連続して受信する通信とを、切替装置の様な特別のハードウェアを用いることなく同一の管理端末装置１０４で行うことが可能である。

【０２３０】また、本実施形態の並列計算機システムの管理装置によれば、特定のノードのメインプロセッサ２０２またはサブプロセッサ２１２が動作時に出力するノードメッセージ４０３を蓄積し、管理端末装置１０４が前記の蓄積されたノードメッセージ４０３を読み取るので、並列計算機システムを構成する複数のノード１００ａ～１００ｃのメインプロセッサ２０２がノードメッセージ４０３を出力した後にその動作を停止した場合であっても、ノードメッセージ４０３を管理端末装置１０４で一括して管理することが可能である。

【０２３１】（実施形態４）以下に、本発明の並列計算機システムの管理装置において、ノード１００ａ～１００ｃのメインプロセッサ２０２のシステム立ち上げ処理を実行し、必要に応じてそのブートストラップデバイスを変更してシステム立ち上げ処理を行う実施形態４について説明する。

【０２３２】本実施形態の並列計算機システムの管理装置では、ブートストラップデバイスからのロード処理が異常終了した場合に、ノード１００ａ～１００ｃのメインメモリ２０４の内容を読み書きすることによってその内容を変更し、ノード１００ａ～１００ｃのメインプロセッサ２０２をリセットすることによって、他のブートストラップデバイスからのロード処理を行うことが可能である。

【０２３３】本実施形態の並列計算機システムの管理装置において、ノード１００ａ～１００ｃのメインメモリ２０４の内容を読み書きする場合には、「ＭＳ－ＲＥＡＤ」コマンド及び「ＭＳ－ＷＲＩＴＥ」コマンドを使用する。これらのシステム制御コマンドは、ノード１００ａ～１００ｃのメインプロセッサ２０２にて通常の業務として並列処理を実行中に障害が発生したときに、ノード１００ａ～１００ｃのメインメモリ２０４の内容を調査する場合にも使用することが可能である。

【０２３４】また、本実施形態の並列計算機システムの管理装置では、ノード１００ａ～１００ｃのメインプロセッサ２０２、サブプロセッサ２１２及びプロセッサメモリ制御機構２０５等のハードウェアモジュール内のレ

ジスタの内容を読み書きすることも可能であり、その場合には、「ＲＥＧ－ＲＥＡＤ」コマンド及び「ＲＥＧ－ＷＲＩＴＥ」コマンドを使用する。

【０２３５】例えば、ノード１００ａ～１００ｃのメインプロセッサ２０２にて、通常の業務である並列処理を実行中に障害が発生した場合は、ノード１００ａ～１００ｃ内の各ハードウェアリソースが採取する障害ログをレジスタに退避しておき、前記の「ＲＥＧ－ＲＥＡＤ」コマンドにより管理端末装置１０４から前記障害ログを読み出すことにより、管理端末装置１０４からの障害要因の特定が可能となる。

【０２３６】また、本実施形態の並列計算機システムの管理装置では、「ＳＴＡＴＵＳ－ＲＥＡＤ」コマンドにより、ノード１００ａ～１００ｃのステータスコードを読み出し、システムダウンを起こしているノードがあれば、「ＰＲＯＣ－ＲＥＳＥＴ」コマンドを送信することで、前記のシステムダウンを起こしているノードのメインプロセッサ２０２をリセットし、再起動させるオペレーションも可能となる。

【０２３７】以下に、本実施形態の並列計算機システムの管理装置におけるノード１００ａ～１００ｃのメインプロセッサ２０２のシステム立ち上げ処理について説明する。

【０２３８】図１３は、本実施形態の並列計算機システムの管理装置におけるノード１００ａ～１００ｃのメインプロセッサ２０２のシステム立ち上げ処理手順を示すフローチャートである。

【０２３９】図１４は、本実施形態の並列計算機システムの管理装置におけるノード１００ａ～１００ｃのＳＲＡＭ２１４のメモリマップを示す図である。図１４において、１４００はプライマリブートストラップパス情報、１４０１はオルタネートブートストラップパス情報である。

【０２４０】図１４に示す様に、本実施形態の並列計算機システムの管理装置におけるノード１００ａ～１００ｃのＳＲＡＭ２１４のメモリマップは、オペレーティングシステム等のソフトウェア２０３を格納している第１のブートストラップデバイスを示すプライマリブートストラップパス情報１４００と、第１のブートストラップデバイスが使用できない場合に使用するブートストラップデバイスを示すオルタネートブートストラップパス情報１４０１とを備えている。

【０２４１】図１３に示す様に、本実施形態の並列計算機システムの管理装置におけるノード１００ａ～１００ｃのメインプロセッサ２０２のシステム立ち上げ処理手順では、管理端末装置１０４からの電源投入指示により主電源２００が投入されると、メインプロセッサ２０２によりブートストラップＲＯＭ２１０に格納されているブートストラッププログラムが実行され、ステップ１３００の処理にて、ＳＲＡＭ２１４内のパネルステータス管理領域のステータスコードを「１０００」とし、ステップ１３０１の処理にて、ノード１００ａ～１００ｃ内の各ハードウェアモジュールの初期診断及び初期化を行う。

【０２４２】ステップ１３０２の処理では、ステップ１３０１の処理のハードウェアの初期診断及び初期化が正常終了したかどうかをチェックし、ステップ１３０１の処理でハードウェアの初期診断及び初期化が正常終了している場合には、ステップ１３０３の処理に進む。

【０２４３】ステップ１３０１の処理でハードウェアの初期診断及び初期化が異常終了している場合には、ステップ１３１３の処理にて、当該ノードに備えられたパネルにステータスコード「１ＦＦＦ」を表示し、当該ノードのメインプロセッサ２０２のシステム立ち上げ処理は異常終了となる。

【０２４４】ステップ１３０３の処理では、ステータスコードを「２０００」とし、ステップ１３０４の処理にて、ＳＲＡＭ２１４に格納されているハードウェア依存情報のうち、図１４に示すブートストラップパス情報を参照し、プライマリブートストラップパス情報１４００にて指定されるブートストラップデバイス（例えば、システムディスク２０７等）から、オペレーティングシステム等のソフトウェア２０３をメインメモリ２０４にロードする。

【０２４５】ＳＲＡＭ２１４内のブートストラップパス情報は、ブートストラップＲＯＭ２１０に格納されているブートストラッププログラムの実行時にメインメモリ２０４内の特定領域にコピーされ、システムが立ち上がるとソフトウェア２０３にて参照可能となる。

【０２４６】また、本実施形態の並列計算機システムの管理装置では、ブートストラップデバイスには、自ノード内のローカルファイルの他にイーサネット経由（システム制御インタフェース）にてｂｏｏｔｐプロトコル（Ｒｅｑｕｅｓｔ　Ｆｏｒ　Ｃｏｎｎｅｃｔブートのベースとなるプロトコル）を使用し、イーサネットに接続される他のノードから取得可能となるブートストラップファイルも適用可能である。

【０２４７】ステップ１３０５の処理では、プライマリブートストラップパス情報１４００にて指定されるブートストラップデバイスからオペレーティングシステム等のソフトウェア２０３をメインメモリ２０４にロードするロード処理が正常終了したかどうかをチェックしており、前記のロード処理に成功すると、ステップ１３０６の処理に進み、失敗するとステップ１３１４の処理に進む。

【０２４８】ステップ１３０６の処理にて、ステータスコードを「３０００」とし、メインメモリ２０４にロードされたソフトウェア２０３が起動され、ステップ１３０７の処理でステータスコードを「Ａ０００」とし、ステップ１３０８の処理にて各種システムパラメータを設

定し、ステップ１３０９の処理にて、ファイルシステムの初期化を行い、ステップ１３１０の処理にて、ＴＣＰ／ＩＰなどのネットワークの初期化を行う。

【０２４９】本実施形態の並列計算機システムの管理装置では、ノード１００ａ～１００ｃのメインプロセッサ２０２で動作するオペレーティングシステム及びネットワークソフトウェア等のソフトウェア２０３の機能を使用するシステム運用支援インタフェースは、この時点で使用可能となる。

【０２５０】ステップ１３１１の処理にて、アプリケーションソフトウェアの起動を行い、ステップ１３１２の処理にてステータスコードを「Ｆ０００」とし、メインプロセッサ２０２のシステム立ち上げ処理を終了する。

【０２５１】一方、ステップ１３１４の処理では、ＳＲＡＭ２１４内のオルタネートブートストラップパス情報１４０１を参照し、オルタネートブートストラップパス情報１４０１にて指定されるブートストラップデバイス（本実施形態の並列計算機システムの管理装置では特に開示していないが、ＤＡＴ（Ｄｉｇｉｔａｌ　Ａｕｄｉｏ　Ｔａｐｅ）等の入出力装置）からのオペレーティングシステム等のソフトウェア２０３をメインメモリ２０４にロードする。

【０２５２】ステップ１３１５の処理にて、オルタネートブートストラップパス情報１４０１にて指定されるブートストラップデバイスからのロードに成功したかどうかをチェックし、成功するとステップ１３０６の処理に進む。

【０２５３】ステップ１３１５の処理にて、オルタネートブートストラップパス情報１４０１にて指定されるブートストラップデバイスからのロードが成功しない場合、ステップ１３１６の処理にて、オペレータによるブートストラップデバイス指定によりロード処理を行う。

【０２５４】ステップ１３１７の処理にて、ステップ１３１６の処理でのオペレータのブートストラップデバイス指定によるロード処理が正常終了したかどうかをチェックし、正常終了している場合にはステップ１３０６の処理に進み、正常終了していない場合には、ステップ１３１８の処理にて、ステータスコードを「２ＦＦＦ」とし、メインプロセッサ２０２のシステム立ち上げ処理が異常終了する。

【０２５５】前記の様にして行ったノード１００ａ～１００ｃのメインプロセッサ２０２のシステム立ち上げ処理が異常終了した場合には、さらに、以下の様に、ブートストラップデバイスを変更したシステム立ち上げ処理を行う。

【０２５６】管理端末装置１０４のソフトウェア３０１は、「ＭＳ－ＲＥＡＤ」コマンドを使用して、システム制御インタフェース経由にて、ノード１００ａ～１００ｃのメインメモリ２０４のブートストラップパス情報が格納されている前記特定領域を参照し、メインプロセッサ２０２のシステム立ち上げ処理に失敗したブートストラップデバイスを確認する。

【０２５７】次に、管理端末装置１０４のソフトウェア３０１は、「ＭＳ－ＷＲＩＴＥ」コマンドを使用し、システム制御インタフェース経由にて、ノード１００ａ～１００ｃのメインメモリ２０４のブートストラップパス情報が格納されている前記特定領域に、メインプロセッサ２０２のシステム立ち上げ処理に失敗したブートストラップデバイス以外のブートストラップデバイス名を書き込む。

【０２５８】管理端末装置１０４のソフトウェア３０１は、前記の様に、ノード１００ａ～１００ｃのメインメモリ２０４の前記特定領域のブートストラップパス情報を書き替えた後、「ＰＲＯＣ－ＲＥＳＥＴ」コマンドを使用し、ノード１００ａ～１００ｃのメインプロセッサ２０２をリセットしてメインプロセッサ２０２のシステム立ち上げ処理を再度行うことで、ブートストラップ先を変更したシステム立ち上げ処理を行うことが出来る。

【０２５９】また、ブートストラップパス情報の書き換えについては、ノード１００ａ～１００ｃのメインプロセッサ２０２のシステム立ち上げ処理が正常終了している場合には、以下の方法でも可能である。

【０２６０】すなわち、ノード１００ａ～１００ｃのＳＲＡＭ２１４のブートストラップパス情報は、ノード１００ａ～１００ｃのソフトウェア２０３からも書き換え可能であるので、管理端末装置１０４のソフトウェア３０１は、システム運用支援インタフェース経由にて、ノード１００ａ～１００ｃのソフトウェア２０３に対し、ブートストラップパス情報の書き換えを指示し、指示されたソフトウェア２０３が当該ノードのブートストラップパス情報を書き替える。

【０２６１】ノード１００ａ～１００ｃのソフトウェア２０３は、更新されたブートストラップパス情報をシステム制御インタフェース経由にて管理端末装置１０４のソフトウェア３０１に通知し、管理端末装置１０４のソフトウェア３０１が、システム制御インタフェース経由にて、前記「ＰＲＯＣ－ＲＥＳＥＴ」コマンドを使用してノード１００ａ～１００ｃのメインプロセッサ２０２をリセットすれば、直ちに更新されたブートストラップパスからのロード処理が行われる。

【０２６２】以上説明した様に、本実施形態の並列計算機システムの管理装置によれば、管理端末装置１０４からの指示によりノード１００ａ～１００ｃのメインメモリ２０４またはレジスタの内容を参照または更新するので、並列計算機システムを構成する複数のノード１００ａ～１００ｃの障害発生時のメインメモリ２０４の内容を管理端末装置１０４で一括して管理することが可能である。

【０２６３】また、本実施形態の並列計算機システムの管理装置によれば、管理端末装置１０４からの指示によ

りノード１００ａ～１００ｃのメインプロセッサ２０２のリセットを行うので、並列計算機システムを構成する複数のノード１００ａ～１００ｃのメインプロセッサ２０２のリセットを管理端末装置１０４から一括して行うことが可能である。

【０２６４】また、本実施形態の並列計算機システムの管理装置によれば、管理端末装置１０４は、ノード１００ａ～１００ｃとの間のインタフェースを使い分けることが可能であり、管理端末装置１０４からの指示によりノード１００ａ～１００ｃのメインメモリ２０４のブートストラップパス情報を変更し、メインプロセッサ２０２のリセットを行うので、並列計算機システムを構成する複数のノード１００ａ～１００ｃの特定のブートストラップデバイスに障害が発生した場合に、管理端末装置１０４からの指示により、ブートストラップデバイスを変更してノード１００ａ～１００ｃのメインプロセッサ２０２のシステム立ち上げ処理を行うことが可能である。

【０２６５】（実施形態５）以下に、本発明の並列計算機システムの管理装置において、複数の管理端末装置を用いて信頼性を向上させた実施形態５の概略構成について説明する。

【０２６６】図１５は、本発明の並列計算機システムの管理装置において、管理端末装置を二重化した実施形態５の概略構成を示す図である。図１５において、１０５ｅはシステム制御機構、１０６ｅは通信ケーブル、１０８ｅはＬＡＮ制御機構、１０９ｅは通信ケーブル、１１１は管理端末装置である。

【０２６７】図１５に示す様に、本実施形態の並列計算機システムの管理装置は、管理端末装置１１１と、通信ケーブル１０６ｅと、通信ケーブル１０９ｅとを備え、管理端末装置１１１は、システム制御機構１０５ｅと、ＬＡＮ制御機構１０８ｅとを有しており、管理端末装置１１１のシステム制御機構１０５ｅを通信ケーブル１０６ｅを介してネットワーク集線装置１０７に接続し、管理端末装置１１１のＬＡＮ制御機構１０８ｅを通信ケーブル１０９ｅを介してネットワーク集線装置１１０に接続している。

【０２６８】前記の様に、本実施形態の並列計算機システムの管理装置では、複数の管理端末装置１０４及び１１１を備えているので、１つの管理端末装置が故障しても、他の管理端末装置により、並列計算機システムの運用管理を続行することが可能であるが、複数の管理端末装置を同時に使用して並列計算機システムの運用管理を行うと、複数の管理端末装置が送信するシステム制御コマンドやアダプタ制御コマンドの内容が互いに競合することがあるので、複数の管理端末装置を用いているときに管理端末装置の動作の競合を防止する処理が必要になる。

【０２６９】以下に、本実施形態の並列計算機システムの管理装置において複数の管理端末装置を用いているときに管理端末装置の動作の競合を防止する処理手順について説明する。

【０２７０】図１６は、本実施形態の並列計算機システムの管理装置において複数の管理端末装置の動作の競合を防止する処理手順を示すフローチャートである。

【０２７１】本実施形態の並列計算機システムの管理装置において、管理端末装置を二重化している場合には、管理端末装置の二重化情報を、例えば、管理端末装置１０４及び管理端末装置１１１の両方のソフトウェア３０１から参照可能な記憶領域に予め設定しておくことで、二重化した管理端末装置の競合を防止することが可能となる。

【０２７２】図１６に示す様に、本実施形態の並列計算機システムの管理装置において管理端末装置を二重化しているときの処理手順では、ステップ１６００の処理で、管理端末装置１０４及び管理端末装置１１１の両方のソフトウェア３０１は、管理端末装置が二重化されていることを示す二重化ビットを参照し、ビットが立っている場合には、管理端末装置が二重化されていることを認識する。

【０２７３】ステップ１６０１の処理では、ネットワーク（例えば、システム運用支援インタフェース）経由にて、相手の管理端末装置のＩＰアドレスを取得する。

【０２７４】ステップ１６０２の処理では、メイン管理端末装置と、前記メイン管理端末装置をバックアップするサブ管理端末装置とを決定するため、例えば、ＩＰアドレスの若い方をメイン管理端末装置、そうでない方をサブ管理端末装置とする。

【０２７５】このとき、メイン管理端末装置のみを動作させておき、前記メイン管理端末装置に障害が発生したときに、直ちにサブ管理端末装置に切り替える運用方法と、メイン管理端末装置とサブ管理端末装置とを同時に動作させる運用方法とを行うことが可能であるが、後者の場合は、双方からのノード１００ａ～１００ｃを制御するシステム制御コマンドや、アダプタ制御コマンドの内容が競合することがあるため、サブ管理端末装置から送信可能なシステム制御コマンド及びアダプタ制御コマンドを一部制限する。

【０２７６】例えば、ステップ１６０３の処理にて、自管理端末装置がメイン管理端末装置であるかどうかを判定し、メイン管理端末装置でなかった場合には、ステップ１６０４の処理にて、システム制御コマンド（「Ｐ－ＯＮ」「Ｐ－ＯＦＦ」等）や、また、アダプタ制御コマンド（「ＳＥＴ－ＣＯＮＮＥＣＴ」等）を発行禁止にすることで、ノード１００ａ～１００ｃを制御するシステム制御コマンドや、アダプタ制御コマンドの内容が競合しても、並列計算機システムとしての整合性を保つことが可能である。

【０２７７】以上説明した様に、本実施形態の並列計算

機システムの管理装置によれば、複数の管理端末装置を備えているので、１つの管理端末装置に障害が発生した場合でも並列計算機システムの運用管理を続行し、並列計算機システムの信頼性を向上させることが可能である。

【０２７８】また、本実施形態の並列計算機システムの管理装置によれば、複数の管理端末装置にメイン管理端末装置とサブ管理端末装置とを設定するので、並列計算機システムを複数の管理端末装置で管理した場合に、前記複数の管理端末装置の動作の競合を防止することが可能である。

【０２７９】（実施形態６）以下に、本発明の並列計算機システムの管理装置において、管理端末装置１０４に補助電源で動作する電源投入論理を付加し、管理端末装置１０４の主電源を遠隔地から投入することにより並列計算機システムの主電源の投入を行う実施形態６について説明する。

【０２８０】図１７は、本実施形態の並列計算機システムの管理装置における管理端末装置１０４に補助電源で動作する電源投入論理を付加した場合の管理端末装置内のハードウェアの概略構成を示す図である。図１７において、１７００は補助電源、１７０１は電源投入論理、１７０２は電源制御信号、１７０３は主電源、１７０４は端末装置、１７０５はネットワークである。

【０２８１】図１７に示す様に、本実施形態の並列計算機システムの管理装置における管理端末装置１０４は、補助電源１７００と、電源投入論理１７０１と、主電源１７０３とを備え、補助電源１７００から電力の供給を受けている電源投入論理１７０１を電源制御信号１７０２を介して主電源１７０３に接続すると共にネットワーク１７０５を介して別の端末装置１７０４に接続している。

【０２８２】図１７に示す様に、本実施形態の並列計算機システムの管理装置における管理端末装置１０４は、補助電源１７００で動作する電源投入論理１７０１を設けており、電源投入論理１７０１は、ここでは特に図示していないが、ネットワーク制御部、電源制御部及びマイクロプロセッサ等から構成されており、ネットワーク１７０５経由で電源制御指示を受け取ると、主電源１７０３を制御する論理回路を備えている。

【０２８３】この電源投入論理１７０１により、例えば下記のような管理端末装置１０４の遠隔オペレーションが可能となる。

【０２８４】本実施形態の並列計算機システムの管理装置において、ネットワーク１７０５で接続された別の端末装置１７０４は、例えばｔｅｌｎｅｔプロトコルを使用して、電源投入論理１７０１にログインする。（この時、管理端末装置１０４には補助電源１７００が投入されている状態である。）

次に、端末装置１７０４は、電源投入論理１７０１にパワーオンコマンドを発行する。電源投入論理１７０１は、パワーオンコマンドを受け取ると、外部から電源投入指示があったことを認識し、電源制御信号１７０２を出力し、管理端末装置１０４の主電源１７０３を投入する。

【０２８５】管理端末装置１０４の主電源１７０３が投入されると、ブートストラップＲＯＭ３０３に格納されているブートストラッププログラムが管理端末装置１０４のシステム立ち上げ処理を行い、ソフトウェア３０１を起動する。

【０２８６】図１８は、本実施形態の並列計算機システムの管理装置における管理端末装置１０４のシェルプログラムの一例を示す図である。ここで、シェルプログラムとは、汎用のオペレーティングシステムであるＵＮＩＸで実行される複数のコマンド名またはプログラム名を記載した、一連の手続きを行うプログラムを指すが、図１８においては、ＵＮＩＸのコマンド名またはプログラム名の代わりに、そのコマンドの機能を簡単に記載している。

【０２８７】図１８に示す様に、本実施形態の並列計算機システムの管理装置の管理端末装置１０４のシステム立ち上げ処理の際に実行されるシェルプログラムに、予め、ノード１００ａ～１００ｃの主電源２００を投入指示するシステム制御コマンドを記載しておき、管理端末装置１０４の主電源１７０３が投入されたときに、このシェルプログラムが実行されるようにしておく。

【０２８８】このようにすることで管理端末装置１０４の起動を契機として、ノード１００ａ～１００ｃの主電源２００を投入し、ノード１００ａ～１００ｃのメインプロセッサ２０２のシステム立ち上げ処理を自動的に行うことが可能である。

【０２８９】以上説明した様に、本実施形態の並列計算機システムの管理装置によれば、遠隔地からのアクセスにより管理端末装置１０４の主電源１７０３を投入し、さらにノード１００ａ～１００ｃの起動（主電源２００の投入）が可能となり、並列計算機システムの運用管理を遠隔地から行うことができる。

【０２９０】以上、説明してきた本実施形態の並列計算機システムの管理装置では、特に図示していないが、下記のようなシステムにも適用可能である。

【０２９１】（１）各ノードに汎用のオペレーティングシステムを搭載していない、特定の機能を実行する専用の並列計算機システムにおいては、汎用のオペレーティングシステムのネットワーク機能を使用しない前記システム制御インタフェースのみを用いて運用管理を行う。

【０２９２】本発明の並列計算機システムの管理装置によれば、前記システム制御インタフェースは、運用管理の対象となるプロセッサとは独立した補助電源とネットワーク機能を備えており、汎用のオペレーティングシステムのＴＣＰ／ＩＰ等のネットワーク機能を使用しない

ので、前記汎用のオペレーティングシステムを搭載していない専用の並列計算機システムにおいても適用することが可能である。

【０２９３】（２）各ノードに補助電源で動作する機能を持たない、或いは補助電源で動作する機能が限定されている様な並列計算機システムにおいては、補助電源を使用しない前記システム運用支援インタフェースのみを用いて管理を行う。

【０２９４】この場合には、主電源の投入等、補助電源を必須とする機能を除き、システム制御インタフェースの機能をシステム運用支援インタフェースによって代行することにより、本発明の並列計算機システムの管理装置を適用することが可能である。

【０２９５】以上、本発明を、前記実施形態に基づき具体的に説明したが、本発明は、前記実施形態に限定されるものではなく、その要旨を逸脱しない範囲において種々変更可能であることは勿論である。

【０２９６】

【発明の効果】本願において開示される発明のうち代表的なものによって得られる効果を簡単に説明すれば、下記のとおりである。

【０２９７】（１）複数のノードの補助電源で動作し、メインプロセッサが使用するネットワークソフトウェア及び通信ケーブルとは別のネットワークソフトウェア及び通信ケーブルを使用して管理端末装置と通信を行うシステム制御機構に対し、前記管理端末装置からシステム制御コマンドを送信し、前記システム制御コマンドを前記補助電源で動作するサブプロセッサで実行することより複数のノードのメインプロセッサの制御を行うので、並列処理を実行するメインプロセッサの動作並びに前記メインプロセッサのオペレーティングシステム及びネットワークソフトウェアの動作とは無関係に、並列計算機システムを構成する複数のノードの運用管理を管理端末装置で一括して行うことが可能である。

【０２９８】（２）管理端末装置からの指示により複数のノードの主電源の投入または切断を行うので、並列計算機システムを構成する複数のノードの電源の投入または切断を管理端末装置で一括または個別に行うことが可能である。

【０２９９】（３）複数のノードへの主電源の投入指示を、予め設定された特定の時間間隔で行うので、並列計算機システムに電力を供給する電源設備の突入電流を低く抑えることが可能である。

【０３００】（４）管理端末装置からの特定のシステム制御コマンドに対する正常なレスポンスが一定時間中に受信されるかどうかを調べるので、並列計算機システムを構成する複数のノードが正常に動作中であるかを管理端末装置で監視することが可能である。

【０３０１】（５）特定のノードのメインプロセッサまたはサブプロセッサが動作時に出力するノードメッセージを蓄積し、管理端末装置が前記の蓄積されたノードメッセージを読み取るので、並列計算機システムを構成する複数のノードのメインプロセッサがノードメッセージを出力した後にその動作を停止した場合であっても、前記ノードメッセージを管理端末装置で一括して管理することが可能である。

【０３０２】（６）管理端末装置からの指示によりノードのメインメモリまたはレジスタの内容を参照または更新するので、並列計算機システムを構成する複数のノードの障害発生時のメインメモリ及びレジスタの内容を管理端末装置で一括して管理することが可能である。

【０３０３】（７）管理端末装置からの指示により複数のノードのメインプロセッサのリセットを行うので、並列計算機システムを構成する複数のノードのメインプロセッサのリセットを管理端末装置から一括して行うことが可能である。

【０３０４】（８）管理端末装置からの指示により複数のノードのメインメモリ中のブートストラップパス情報を変更し、メインプロセッサのリセットを行うので、並列計算機システムを構成する複数のノードの特定のブートストラップデバイスに障害が発生した場合に、管理端末装置からの指示により、ブートストラップデバイスを変更して前記複数のノードのメインプロセッサのシステム立ち上げ処理を行うことが可能である。

【０３０５】（９）複数の管理端末装置を備えることも可能であるので、１つの管理端末装置に障害が発生した場合でも並列計算機システムの運用管理を続行することが可能であり、複数の管理端末装置をメイン管理端末装置とサブ管理端末装置とに設定するので、並列計算機システムを複数の管理端末装置で管理した場合に、前記複数の管理端末装置の動作の競合を防止することが可能である。

【０３０６】（１０）遠隔地からのアクセスにより管理端末装置の主電源を投入し、さらに複数のノードの主電源を投入するので、並列計算機システムの運用管理を遠隔地から行うことが可能である。

【図面の簡単な説明】

【図１】本発明の並列計算機システムの管理装置を実施する実施形態１の概略構成を示す図である。

【図２】実施形態１の並列計算機システムの管理装置において並列計算機システムを構成するノードの概略構成を示す図である。

【図３】実施形態１の並列計算機システムの管理装置における管理端末装置の概略構成を示す図である。

【図４】実施形態１の並列計算機システムの管理装置における管理端末装置と各ノードとの通信シーケンスの一例を示す図である。

【図５】実施形態１の並列計算機システムの管理装置におけるアダプタ制御コマンド及びそのレスポンスのパケットフォーマットを示す図である。

【図６】実施形態１の並列計算機システムの管理装置におけるシステム制御コマンド及びそのレスポンスのパケットフォーマットを示す図である。

【図７】実施形態１の並列計算機システムの管理装置におけるシステム制御機構のモード遷移を示す図である。

【図８】実施形態１の並列計算機システムの管理装置におけるシステム制御機構の非パケットモードでのコネクション状態の遷移を示す図である。

【図９】実施形態１の並列計算機システムの管理装置におけるシステム制御機構のプロセッサの処理手順の一部を示すフローチャートである。

【図１０】実施形態１の並列計算機システムの管理装置におけるシステムサポート機構のサブプロセッサの処理手順の一部を示すフローチャートである。

【図１１】実施形態２の並列計算機システムの管理装置における管理端末装置から各ノードへ主電源の投入を指示する電源投入シーケンスの一例を示す図である。

【図１２】実施形態３の並列計算機システムの管理装置における管理端末装置に各ノードのノードメッセージを表示するシーケンスの一例を示す図である。

【図１３】実施形態４の並列計算機システムの管理装置におけるノードのメインプロセッサのシステム立ち上げ処理手順を示すフローチャートである。

【図１４】実施形態４の並列計算機システムの管理装置におけるノード内のＳＲＡＭ内のメモリマップを示す図である。

【図１５】本発明の並列計算機システムの管理装置において管理端末装置を二重化した実施形態５の概略構成を示す図である。

【図１６】実施形態５の並列計算機システムの管理装置において複数の管理端末装置の動作の競合を防止する処理手順を示すフローチャートである。

【図１７】実施形態６の並列計算機システムの管理装置における管理端末装置に補助電源で動作する電源投入論理を付加した場合の管理端末装置内のハードウェアの概略構成を示す図である。

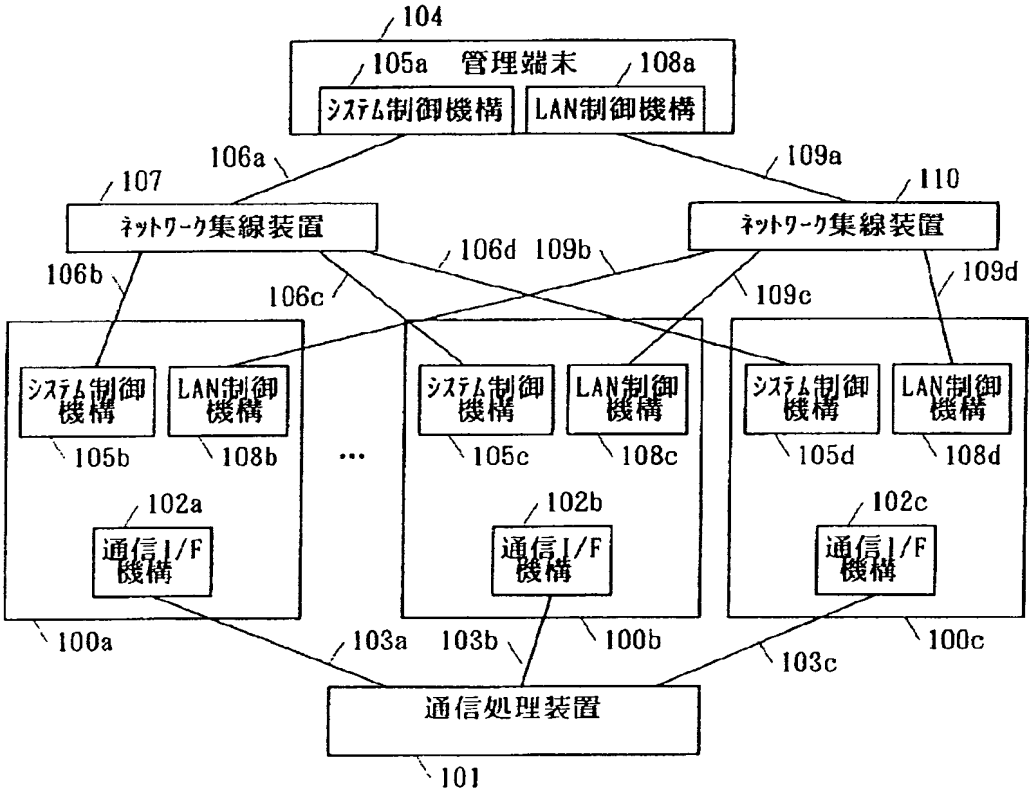
【図１８】実施形態６の並列計算機システムの管理装置における管理端末装置のシェルプログラムを示す。

【符号の説明】

１００ａ～１００ｃ…ノード、１０１…通信処理装置、１０２ａ～１０２ｃ…通信インタフェース機構、１０３ａ～１０３ｃ…通信ケーブル、１０４…管理端末装置、１０５ａ～１０５ｄ…システム制御機構、１０６ａ～１０６ｄ…通信ケーブル、１０７…ネットワーク集線装置、１０８ａ～１０８ｄ…ＬＡＮ制御機構、１０９ａ～１０９ｄ…通信ケーブル、１１０…ネットワーク集線装置、２００…主電源、２０１…補助電源、２０２…メインプロセッサ、２０３…ソフトウェア、２０４…メインメモリ、２０５…プロセッサメモリ制御機構、２０６…システムバス、２０７…システムディスク、２０８…Ｉ／Ｏ制御機構、２０９…ＲＳ－２３２Ｃ制御機構、２１０…ブートストラップＲＯＭ、２１１…システムサポート機構、２１２…サブプロセッサ、２１３…ＲＯＭ、２１４…ＳＲＡＭ、２１５…ローカルバス、２１６…電源投入／切断信号、２１７…プロセッサリセット信号、２１８…ＬＡＮ制御部、２１９…ＲＳ－２３２Ｃ制御部、２２０…プロセッサ、２２１…ＲＯＭ、２２２…ＲＡＭ、２２３…データインタフェース、２２４…制御インタフェース、３００…プロセッサ、３０１…ソフトウェア、３０２…メインメモリ、３０３…ブートストラップＲＯＭ、３０４…プロセッサメモリ制御機構、３０５…システムバス、３０６…Ｉ／Ｏ制御機構、３０７…システムディスク、３０８及び３０９…ＲＳ－２３２Ｃ制御機構、３１０…グラフィックス制御機構、３１１…ＬＡＮ制御部、３１２…ＲＳ－２３２Ｃ制御部、３１３…プロセッサ、３１４…ＲＯＭ、３１５…ＲＡＭ、３１６…制御インタフェース、３１７…データインタフェース、４０１…アダプタ制御コマンド及びそのレスポンス、４０２…システム制御コマンド及びそのレスポンス、４０３…ノードメッセージ、５０１…種別フィールド、５０２…送信元アドレスフィールド、５０３…受信先アドレスフィールド、５０４…情報部フィールド、５０５…識別子、６０１…種別フィールド、６０２…送信元アドレスフィールド、６０３…受信先アドレスフィールド、６０４…情報部フィールド、６０５…識別子、７０１…パケットモード、７０２…非パケットモード、７０３…「ＳＥＴ－ＭＯＤＥ」コマンド、８００…ディスコネクト状態、８０１…ウェイトコネクト状態、８０２…コネクト状態、８０３…「ＳＥＴ－ＣＯＮＮＥＣＴ」コマンド、８０４…システム制御機構間の呼制御、１４００…プライマリブートストラップパス情報、１４０１…オルタネートブートストラップパス情報、１０５ｅ…システム制御機構、１０６ｅ…通信ケーブル、１０８ｅ…ＬＡＮ制御機構、１０９ｅ…通信ケーブル、１１１…管理端末装置、１７００…補助電源、１７０１…電源投入論理、１７０２…電源制御信号、１７０３…主電源、１７０４…端末装置、１７０５…ネットワーク。

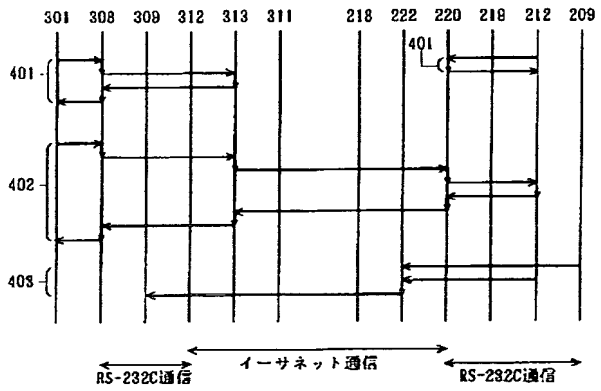
【図１】

# 図１



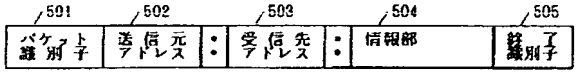

【図４】

# 図４



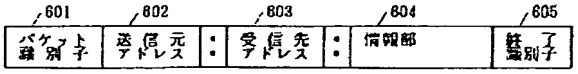

【図５】

# 図５



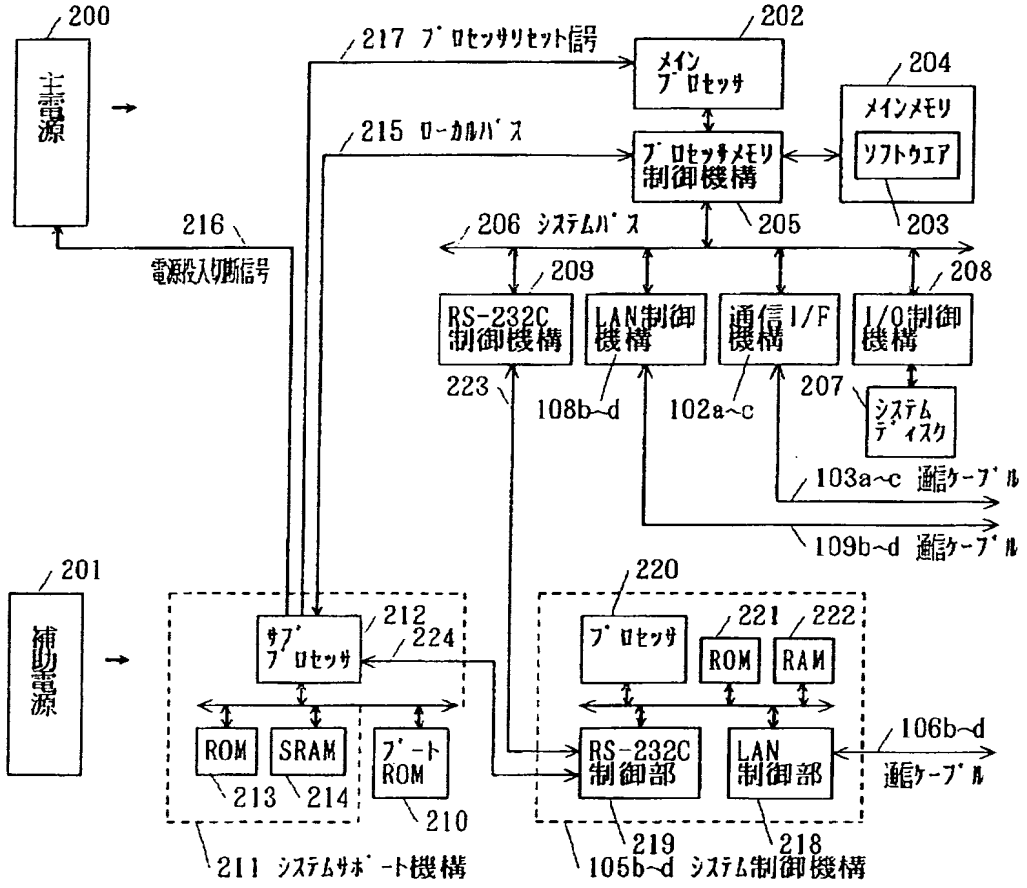

【図６】

# 図６



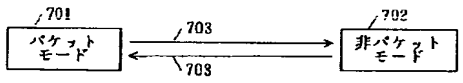
【図２】

図２



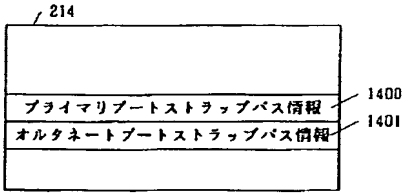

【図７】

図７



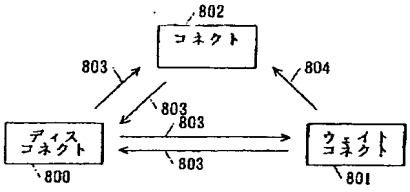

【図１４】

図１４



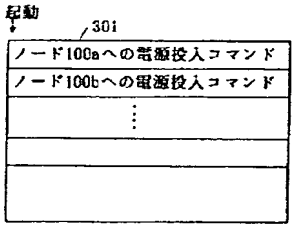

【図８】

図８



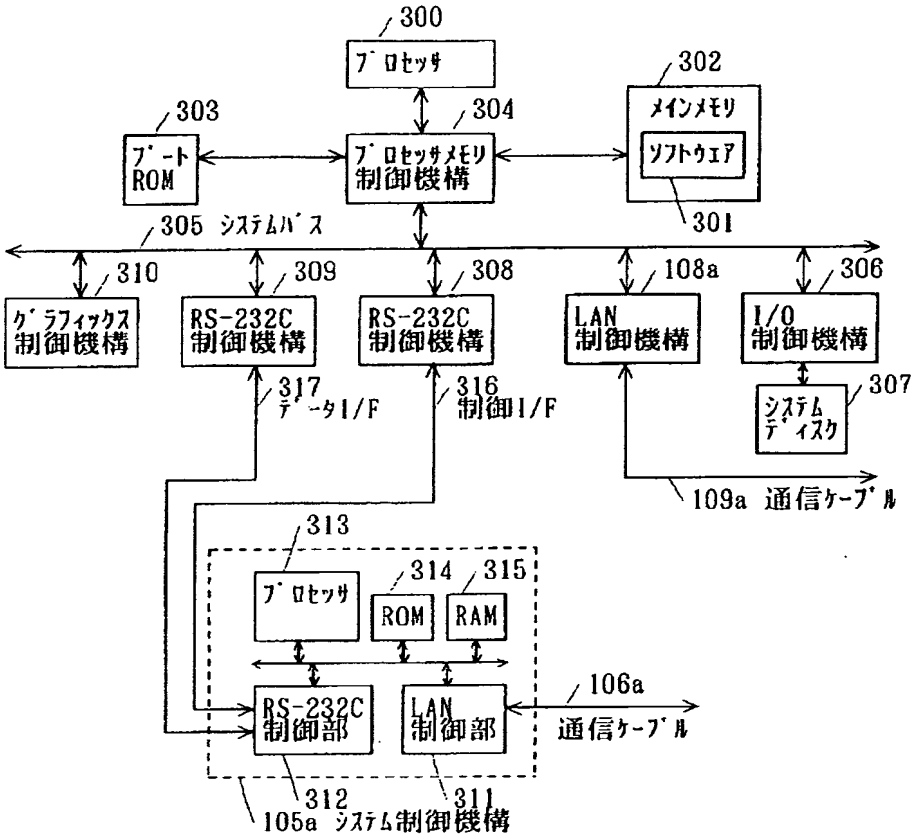

【図１８】

図１８



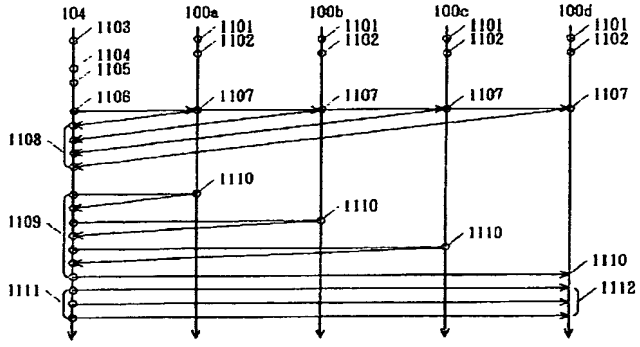
【図３】

# 図３


300
ﾌﾟﾛｾｯｻ
302
ﾒｲﾝﾒﾓﾘ
ｿﾌﾄｳｪｱ
303
ﾌﾞｰﾄ ROM
304
ﾌﾟﾛｾｯｻﾒﾓﾘ 制御機構
301
305 ｼｽﾃﾑﾊﾞｽ
310
ｸﾞﾗﾌｨｯｸｽ 制御機構
309
RS-232C 制御機構
308
RS-232C 制御機構
108a
LAN 制御機構
306
I/O 制御機構
307
ｼｽﾃﾑ ﾃﾞｨｽｸ
317
ﾃﾞｰﾀI/F
316
制御I/F
109a 通信ｹｰﾌﾞﾙ
313
ﾌﾟﾛｾｯｻ
314
ROM
315
RAM
RS-232C 制御部
LAN 制御部
106a
通信ｹｰﾌﾞﾙ
312
311
105a ｼｽﾃﾑ制御機構

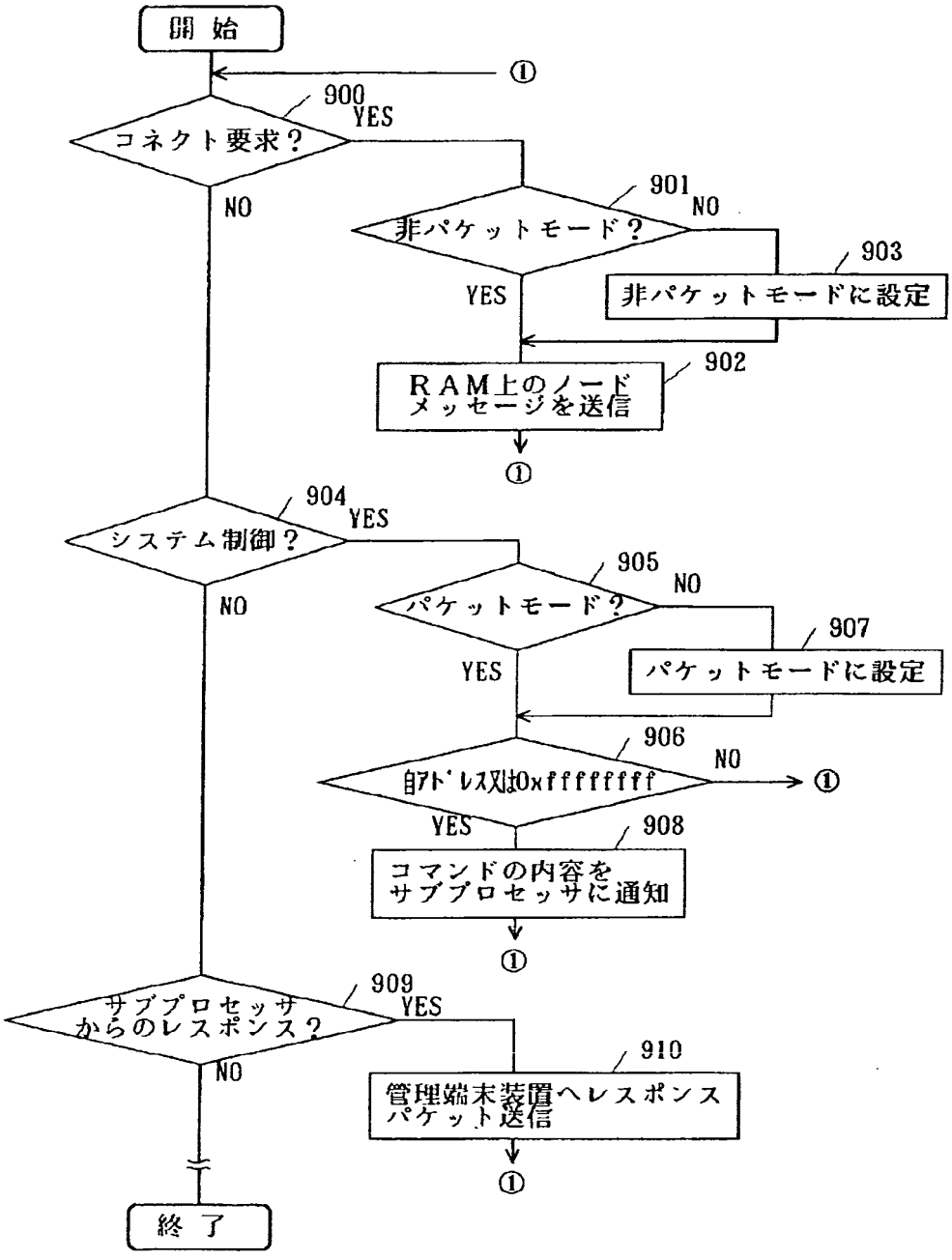
【図１１】

# 図１１


104
100a
100b
100c
100d
1103
1101
1102
1104
1105
1106
1107
1108
1110
1109
1111
1112

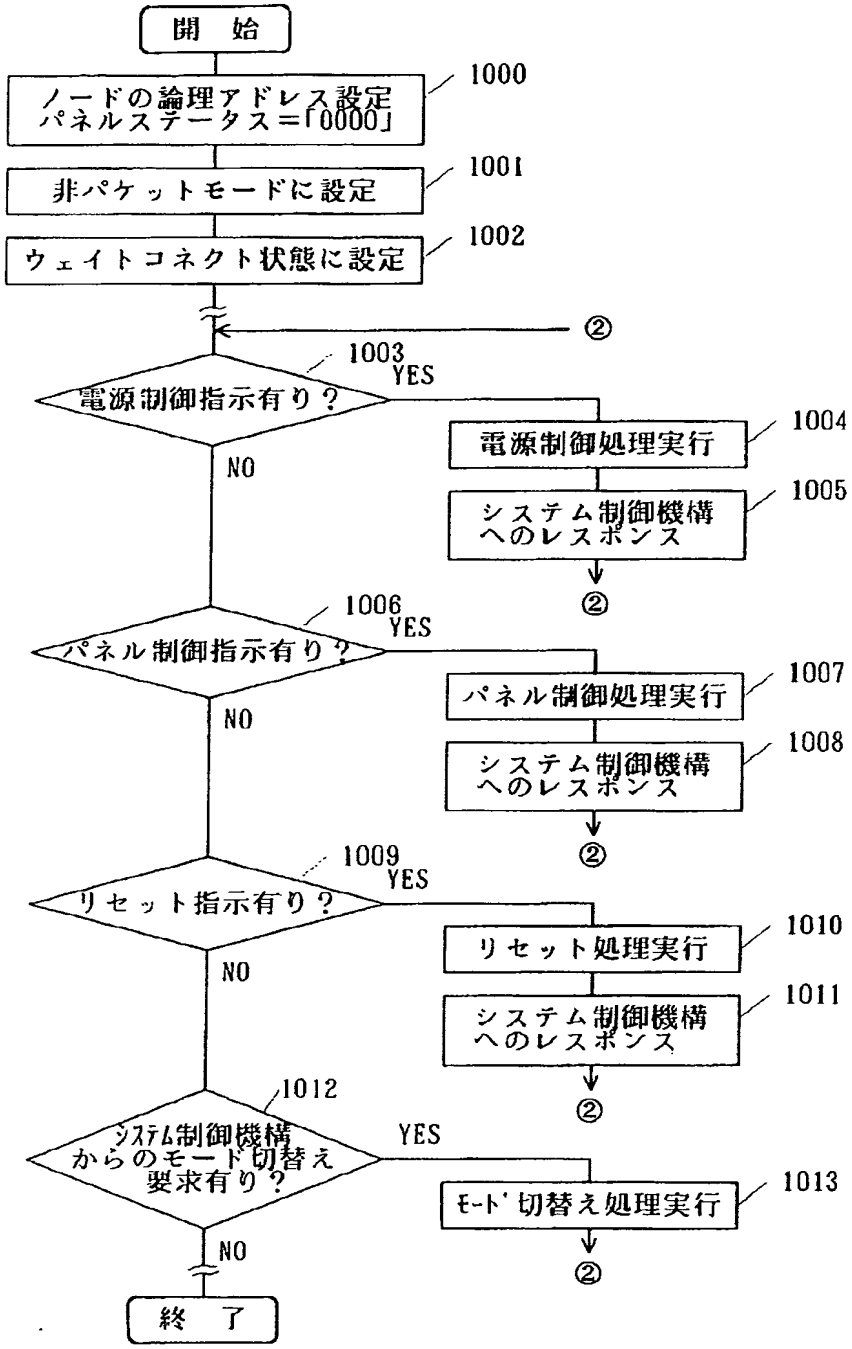
【図９】

# 図９


開 始
①
900
コネクト要求？
YES
NO
901
非パケットモード？
NO
YES
903
非パケットモードに設定
902
ＲＡＭ上のノード
メッセージを送信
①
904
システム制御？
YES
NO
905
パケットモード？
NO
YES
907
パケットモードに設定
906
自ｱﾄﾞﾚｽ又は0xffffffff
NO
①
YES
908
コマンドの内容を
サブプロセッサに通知
①
909
サブプロセッサ
からのレスポンス？
YES
NO
910
管理端末装置へレスポンス
パケット送信
①
終 了

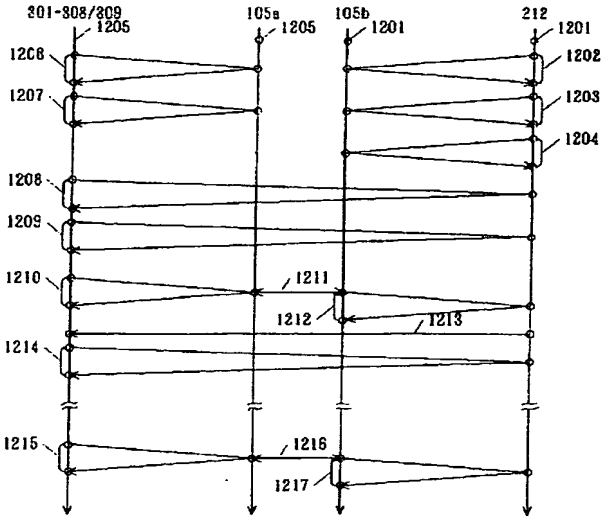
【図１０】

# 図１０


開　始
ノードの論理アドレス設定
パネルステータス＝「0000」
1000
非パケットモードに設定
1001
ウェイトコネクト状態に設定
1002
②
電源制御指示有り？
1003
YES
NO
電源制御処理実行
1004
システム制御機構
へのレスポンス
1005
②
パネル制御指示有り？
1006
YES
NO
パネル制御処理実行
1007
システム制御機構
へのレスポンス
1008
②
リセット指示有り？
1009
YES
NO
リセット処理実行
1010
システム制御機構
へのレスポンス
1011
②
システム制御機構
からのモード切替え
要求有り？
1012
YES
NO
ﾓｰﾄﾞ切替え処理実行
1013
②
終　了

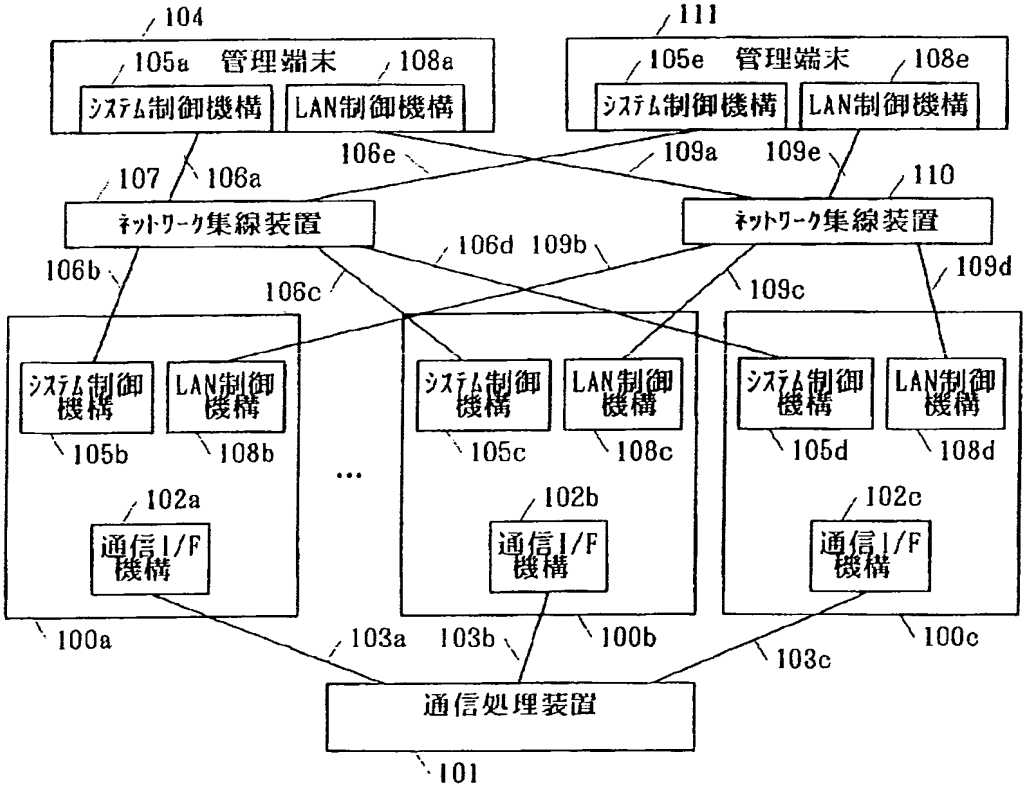
【図１２】

図１２



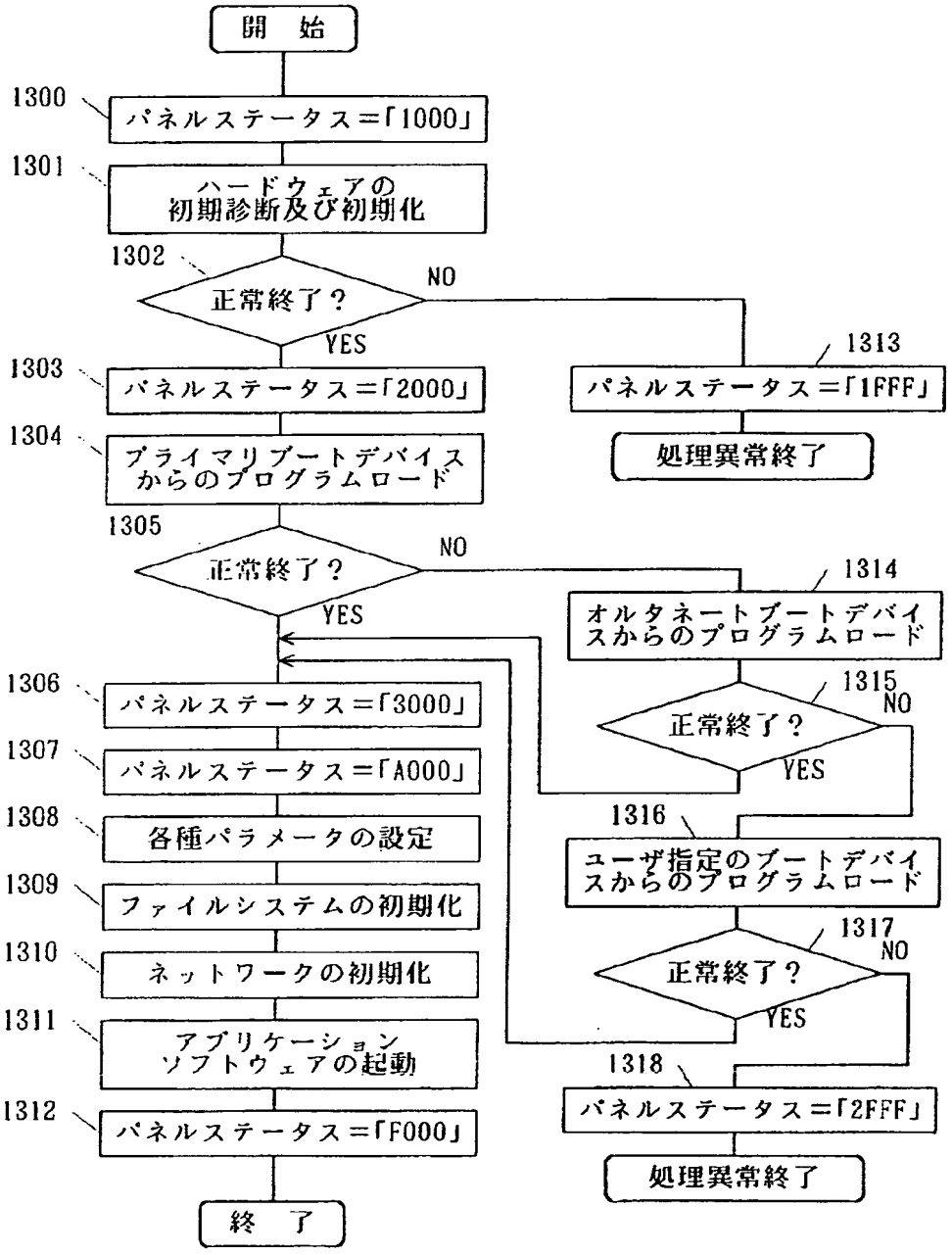

【図１５】

図１５



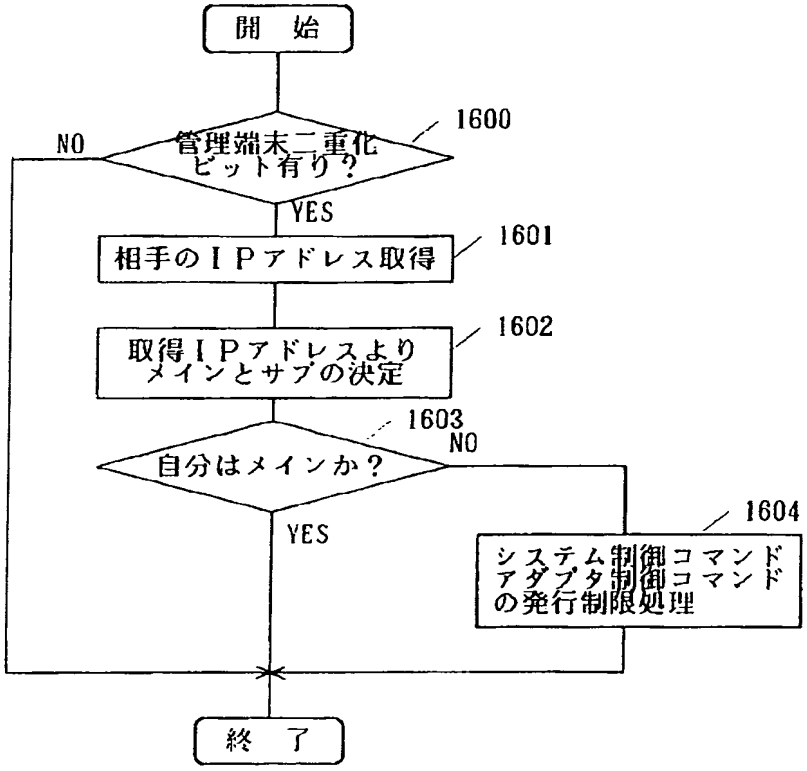
【図１３】

# 図１３


開　始
1300
パネルステータス＝「1000」
1301
ハードウェアの
初期診断及び初期化
1302
正常終了？
NO
YES
1313
パネルステータス＝「1FFF」
処理異常終了
1303
パネルステータス＝「2000」
1304
プライマリブートデバイス
からのプログラムロード
1305
正常終了？
NO
YES
1314
オルタネートブートデバイ
スからのプログラムロード
1315
正常終了？
NO
YES
1306
パネルステータス＝「3000」
1307
パネルステータス＝「A000」
1308
各種パラメータの設定
1309
ファイルシステムの初期化
1310
ネットワークの初期化
1311
アプリケーション
ソフトウェアの起動
1312
パネルステータス＝「F000」
1316
ユーザ指定のブートデバイ
スからのプログラムロード
1317
正常終了？
NO
YES
1318
パネルステータス＝「2FFF」
処理異常終了
終　了

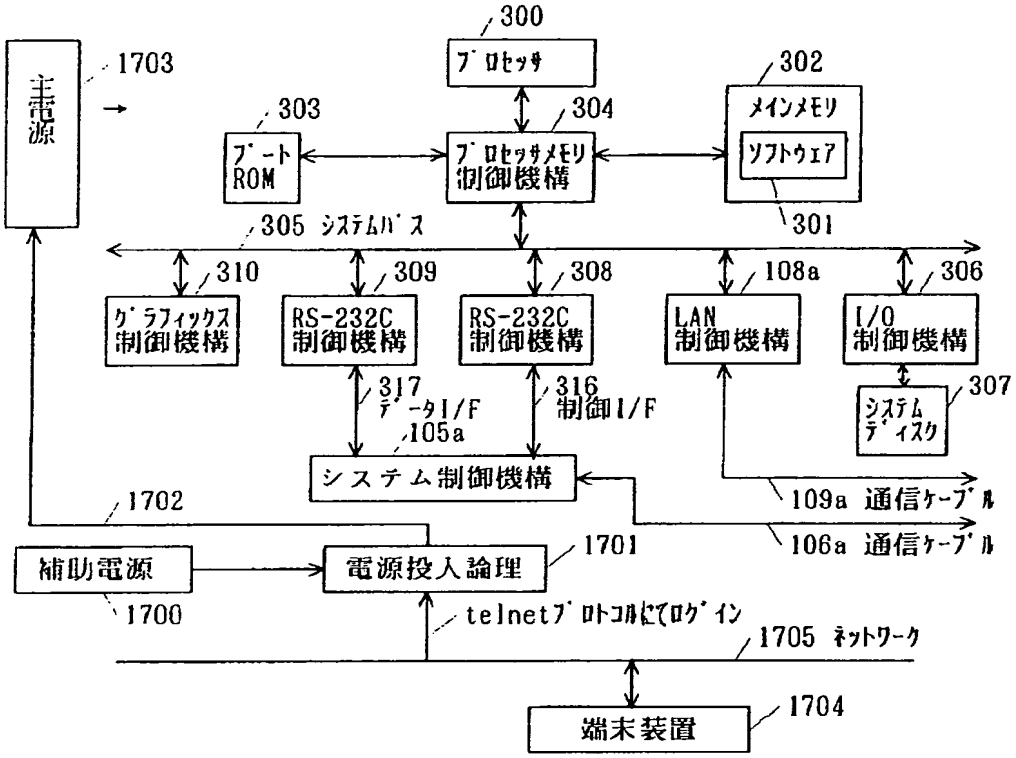
【図１６】

# 図１６


開　始
管理端末二重化ビット有り？
1600
NO
YES
相手のＩＰアドレス取得
1601
取得ＩＰアドレスよりメインとサブの決定
1602
自分はメインか？
1603
NO
YES
システム制御コマンド
アダプタ制御コマンド
の発行制限処理
1604
終　了

【図１７】

# 図１７


主電源
1703
300
ﾌﾟﾛｾｯｻ
302
ﾒｲﾝﾒﾓﾘ
ｿﾌﾄｳｪｱ
303
ﾌﾞｰﾄ ROM
304
ﾌﾟﾛｾｯｻﾒﾓﾘ 制御機構
301
305 ｼｽﾃﾑﾊﾞｽ
310
ｸﾞﾗﾌｨｯｸｽ 制御機構
309
RS-232C 制御機構
308
RS-232C 制御機構
108a
LAN 制御機構
306
I/O 制御機構
307
ｼｽﾃﾑ ﾃﾞｨｽｸ
317 ﾃﾞｰﾀI/F
105a
316 制御I/F
システム制御機構
109a 通信ｹｰﾌﾞﾙ
106a 通信ｹｰﾌﾞﾙ
1702
補助電源
1700
電源投入論理
1701
telnetﾌﾟﾛﾄｺﾙにてﾛｸﾞｲﾝ
1705 ﾈｯﾄﾜｰｸ
端末装置
1704